ESTIMATING END-USE DEMAND: A BAYESIAN APPROACH

Luc Bauwens, Denzil G. Fiebig and Mark F.J. Steel*

Abstract

Eliminating negative end-use or appliance consumption estimates and incorporating direct metering information into the process of generating these estimates; these are two important aspects, of conditional demand analysis (CDA) that will be the focus of this paper. In both cases a Bayesian approach seems a natural way of proceeding. What needs to be investigated is whether it is also a viable and effective approach. In addition, such a framework naturally lends itself to prediction. Our application involves the estimation of electrical appliance consumptions for a sample of Australian households. This application is designed to illustrate the viability of a full Bayesian analysis of the problem.

Key words:
End-use demand, direct metering, non-negative estimates, bayesian conditional demand analysis.

I Introduction

Eliminating negative end-use or appliance consumption estimates and incorporating direct metering information into the process of generating these estimates; these are two important aspects of conditional demand analysis (CDA) that will be the focus of this paper. In both cases a Bayesian approach seems a natural way of proceeding. What needs to be investigated is whether it is also a viable and effective approach.

CDA is designed to disaggregate total energy demand into specific end-use components. For example the household demand for electricity or gas could be disaggregated into demands associated with the different appliances held by the household. Naturally each of these components of total demand should be positive. However, existing CDA studies provide numerous examples where these natural restrictions are violated. Notably, in a study of residential electricity demand the results of Aigner, Sorooshian and Kerwin (1984) included many negative load estimates prompting the comment:

> "While our results show well defined load *shapes* for many appliances, the load *levels* often seem questionable", Aigner et al. (1984, p. 97).

Also in a recent CDA study of residential electricity demand, Fiebig, Bartels and Aigner (1991) used the infrequent occurence (but not elimination) of negative estimates as one measure of the success of their estimators.

CDA relies on heterogeneous household appliance portfolios; essentially end-use demands are estimated by comparing total demand of households with and without a particular appliance. Unfortunately it is sometimes the case that there are high saturation appliances or there is a "bundling" of certain types of appliances each of which gives rise to multicollinearity problems. It is also possible that weak data can arise because of large disturbance variances and/or relatively small sample sizes. Consequently negative estimates in CDA studies can often occur because data are weak or uninformative. One response to this problem is to exploit available prior information; in this case the non-negativity of the end-use demands. We demonstrate how a Bayesian approach can provide a complete solution, guaranteeing nonnegative estimates of the end-use consumptions. Geweke

(1986) has argued forcefully in favour of the Bayesian approach to the problem of inequality constrained linear regression. He contends that it:

"... leads to practical methods for exact inference that are impossible to treat using a sampling-theoretic approach", Geweke (1986, p.128).

One procedure that provides accurate estimates of end-use consumptions and automatically overcomes the nonnegativity problem is to abandon the CDA approach and to attach meters to individual end-uses enabling the direct measurement of the associated consumption. While this is a conceptually straightforward method it is unfortunately not practical because the cost associated with extensive direct metering would be prohibitive. Moreover it ignores the information that is provided by CDA studies using data that is often readily available. The compromise situation where some direct metering information is available and is combined with the CDA information seems to be the appropriate way to proceed.

Typically direct metering information will be available for only a subset of end-uses and/or a small number of households. A natural way to utilize these data is to view it as prior information on the specific appliance consumptions and proceed in a Bayesian framework. Such an approach has been taken by Hsiao, Mountain and Ho (1990) who present three alternative Bayesian methods that share the following general characteristics:

(i) Priors are specified as being multivariate normal.

(ii) Parameters of the priors are "estimated" on the basis of *independent* direct metering information where the direct metering is assumed to come from some previous sample or from a sample of customers that is different from that used in the CDA.

(iii) The posterior distributions of the parameters of interest are characterized solely by the posterior mean. An initial assumption of a known disturbance covariance matrix enables explicit representation of these posterior means as matrix weighted averages of sample and prior information.

Differences in their three methods relate to the generality of the prior distributions chosen. The preferred, and most general method, allows for a structure where the end-use consumptions are themselves related to certain demographic characteristics.

Caves, Herriges, Train and Windle (1987) have also used a Bayesian framework to combine prior information with a CDA model. Here though the prior information is obtained from engineering simulation models. Like Hsiao et al. (1990) they rely on a prior specification that is multivariate normal and the posterior mean is calculated directly as a matrix weighted average of the prior information, provided by engineering data, and the sample information provided by CDA. Apart from the source of the prior information, their analysis differs from Hsiao et al. (1990) in two other important respects. They explicitly recognize the form of the approximations being made to generate the reported point estimates. Specifically the posterior mean is actually that of the distribution that is suggested to approximate the actual posterior distribution. Secondly, the importance of the nonnegativity issue is recognized. As an extension to their main analysis, Caves et al. (1987) admit the inappropriateness of their prior distribution and briefly comment on the use of a truncated normal prior.

There are several important aspects that differentiate our approach from the previous Bayesian analyses of Hsiao et al. (1990) and Caves et al. (1987). In particular:

(i) Unlike Caves et al. the nonnegativity issue is addressed directly by initially incorporating this restriction into our basic specification of the model.

(ii) Allowance is made for different forms and sources of direct metering information. In particular, we explicitly address the situation of having partial direct metering for the same time period and group of households as is to be used for the CDA. The analysis is usefully viewed as a three stage process starting with relatively diffuse priors which are then updated by two sources of data corresponding to direct metering and CDA.

(iii) The underlying data processes associated with the direct metering information and the CDA data are taken to be elliptical and hence are both more general than the assumption of normality.

(iv) A more complete description and analysis of the posterior distribution than simply the calculation of posterior means is provided and no recourse is made to approximations.

(v) The analysis is extended to prediction on the basis of a complete

4

characterization of the predictive density.

Our application involves the estimation of electrical appliance consumptions for a sample of Australian households. This application is designed to illustrate the viability of a full Bayesian analysis of the problem. Inevitably the attraction of these methods to practitioners is related to computational matters. We pay particular attention to this aspect of the problem.

II Conditional demand analysis and direct metering

The CDA model

The basic CDA model can be written as

$$(2.1) \qquad y_i = \sum_{j=1}^{k} UEC_{ij} \; D_{ij} + v_i$$

where

y_i = energy consumption by household i for some period of time (possibly a year, month, day or hour),

D_{ij} = an indicator of ownership of appliance j by household i; D_{i1} is unity for all i to represent end-uses such as lighting that are held by all households,

UEC_{ij} = unit of energy consumption of appliance j by household i,

v_i = a random disturbance.

Both y_i and D_{ij} are observed; the former coming from billing records and the latter from household surveys. It is the end-use energy consumptions (UEC_{ij}) that are typically not observed and need to be determined. Treating the UEC_{ij} as constant over households, as in say Caves et al. (1987), implies that (2.1) becomes a regression on a set of appliance dummy variables. The estimated coefficients of these dummy variables would then be interpreted as estimated UEC_js.

A more general specification allows the end-use consumptions to vary across households by making them depend on a variety of household characteristics. Assuming these relationships are linear we obtain a general CDA model which

we write in matrix form as

$$(2.2) \qquad y_c = Z_c \phi + D_c \gamma + u_c$$
$$\equiv X_c \beta + u_c$$

where y_c and u_c are $N_c \times 1$ vectors and Z_c and D_c are $N_c \times p$ and $N_c \times k$ matrices of regressors with union X_c. D_c is the matrix of observations on the set of k appliance dummies while Z_c contains the set of explanatory variables that determine household appliance utilization.

Following the basic CDA model specification we assume that appliance ownership is exogenous. This is somewhat contentious because it is quite reasonable to argue that some of the factors affecting appliance utilization also affect decisions to acquire appliances. However, in the short run appliance stocks are fixed and existing evidence suggests that the bias from ignoring possible endogeneity may be qualitatively small; see Sebold and Parris (1989).

There are several reasons for dividing the regressors of (2.2) into two groups. Importantly it highlights the parameters of most interest, namely the elements of γ. By carefully defining the individual columns of Z_c a useful interpretation is provided for γ. Because of (2.1) the columns of Z_c will be in the form of interactions of demographic variables with appliance dummies. Consider a particular interaction. The appliance dummy is multiplied by the associated demographic variable which is defined in terms of deviations from its mean where the mean is calculated over those households possessing the appliance in question. The jth element of γ can now be interpreted as the energy consumption of the jth appliance for the "average household" possessing that appliance.

This interpretation highlights the fact that γ should clearly never be negative. Such prior information is important and should be incorporated into the estimation of CDA models such as (2.2).

Notice that in moving from (2.1) to (2.2) the disturbance terms have been differentiated. Typically one would expect a heteroskedastic u_c. In particular it is likely that disturbance variances would be related to

6

appliance holdings.

Direct metering

CDA provides an indirect method of estimating end-use energy consumption. An alternative method of estimating these is to meter specific appliances directly. For example in Australia households can opt for an off-peak hot water system that is charged at a lower than normal rate. Because of its differentiated price, this end-use needs to be metered separately. Alternatively limited direct metering programs have also been conducted as part of the load research programs of utilities. In the U.S.A. there have been several examples of such programs; see Sebold and Parris (1989) for a summary.

Direct metering information may have been collected independently of the CDA data for a different set of households or for a different time period. Alternatively it could have been collected in conjunction with the CDA data. Because of the costs involved in metering individual appliances, the typical situation will be one where there is limited direct metering information covering only a subset of appliances and a sample of households probably smaller than that available for CDA analysis. Moreover, Bartels and Fiebig (1990) argue that if one has the opportunity to choose where to place meters then it is preferable in terms of the precision gains for estimation to spread the meters over a number of appliances rather than to concentrate on a particular appliance.

Following Fiebig et al. (1991) and Hsiao et al. (1990), the direct metering information will be combined with CDA to provide improved estimates of end-use consumptions. The only restriction to be imposed on the direct metering information is that its sampling process should have the same structure as the CDA model of (2.2) with the same coefficients, but possibly different covariance structure. Specifically write the sampling process for the N_d direct metering observations as

$$(2.3) \qquad y_d = Z_d \phi + D_d \gamma + u_d$$
$$\equiv X_d \beta + u_d$$

where y_d and u_d are $N_d \times 1$ vectors and Z_d and D_d are $N_d \times p$ and $N_d \times k$ matrices of

regressors with union X_d. Notice that in order to validate use of the CDA models in (2.2) and (2.3) we require that X is weakly exogenous for inference on β; see Engle et al. (1983).

In the case where both sources of information pertain to the same sample of households for the same period, some care needs to be taken in specifying (2.2). In particular the availability of direct metering of appliance j for household i implies that y_{1c} will refer to total consumption *net* of the direct metering consumption. In addition, the ith rows of Z_c and D_c will be modified so as to indicate the absence of appliance j. This avoids using the same information twice.

III Bayesian conditional demand analysis

Three Bayesian CDA models are developed that differ in the specification of the stochastic component of the model. Each one has distinctive features that make them attractive potential models. In practice this provides a selection of alternatives that should cover a wide range of applications.

The common precision model
As already explained, there are two distinct sources of data information pertaining to our problem and we use the same regression models for both, differing only in the covariance structure. The common precision (CP) model entails the stacking of these data as represented by (2.2) and (2.3) to yield:

$$(3.1) \qquad y_s = X_s \beta + u_s$$

where $y_s' = (y_d' \ y_c')$, $X_s' = (X_d' \ X_c')$ and $u_s' = (u_d' \ u_c')$. The stochastic structure on u_s is taken to be multivariate elliptical (or ellipsoidal). Provided symmetry around zero is a reasonable assumption, this provides considerable flexibility. In particular, we assume

$$(3.2) \qquad p(u_s|\delta_s) = |V_s|^{-1/2} \delta_s^{N/2} g_s(\delta_s u_s' V_s^{-1} u_s)$$

where δ_s is a positive scalar precision parameter assumed to be common to

both types of data, $g_s(.)$ is a known nonnegative function such that $r^{(N-2)/2} g_s(r)$ is integrable in r over $\mathbb{R}_+$, $N = N_d + N_c$ and V_s is a known NxN positive definite symmetric (PDS) matrix. By choosing different functional forms of $g_s(.)$ a large class of distributions can be covered: normal, Student t, Pearson type II etc. See Kelker (1970), Dickey and Chen (1985) and Johnson (1987) for details.

Apart from the possible dependencies allowed for by the assumption of an elliptical distribution for u_s, the CP model can accommodate a range of correlation and heteroskedastic structures through the specification of V_s. However the constraint is that V_s must be known or that we are willing to condition on particular choices for the parameters that describe it. At the cost of increasing computational complexity, V_s can be made to depend on unknown parameters as well as on X_s. For example this would be required for a more formal treatment of heteroskedasticity as in Bauwens and Lubrano (1991).

The CP model is now completed by assuming a prior density on the parameters (β, δ_s). For the precision parameter we shall specify the usual improper prior structure:

$$(3.3) \qquad p(\beta, \delta_s) = p(\delta_s)\, p(\beta) \propto \delta_s^{-1}\, p(\beta)$$

where functional independence between β and the precision parameter is assumed. Such a reference prior is shown by Osiewalski and Steel (1992) to result in exactly the same posterior and predictive results for *any* choice of $g_s(.)$ in (3.2). The analysis is then *fully robust* with respect to any departures from normality in the entire class of multivariate elliptical data densities. These results hold irrespective of the specification of $p(\beta)$ in (3.3).

In this framework it follows directly that the marginal posterior for β is given by:

$$(3.4) \qquad p(\beta \,|\, y_s, X_s) \propto p(\beta)\, f_s^m(\beta \,|\, N - m,\, \hat{\beta}_s,\, s_s^{-2}\, X_s{}' V_s^{-1} X_s)$$

where $m = k + p$, $f_s^r(x|\nu, \mu, H)$ denotes an r-variate Student t density with $\nu > 0$ degrees of freedom, location vector $\mu \in \mathbb{R}^r$ and PDS rxr precision matrix H, and where we use the hyperparameters

(3.5) $\hat{\beta}_s = (X_s'V_s^{-1}X_s)^{-1}X_s'V_s^{-1}y_s$

and

(3.6) $s_s^2 = (y_s - X_s\hat{\beta}_s)' V_s^{-1}(y_s - X_s\hat{\beta}_s)/(N - m)$.

Specifying a conventional (unrestricted uniform) noninformative prior for β means that the posterior for β given in (3.4) is a multivariate Student t density. However the prior on β should reflect the nonnegativity of the direct appliance consumptions in γ which is firmly based on their technical interpretation. Assuming the prior mass is spread evenly over the positive orthant leads to a prior given by

(3.7) $p(\beta, \delta_s) \propto \delta_s^{-1} I_+(\gamma)$

where $I_+(\gamma) = 1$ if $\gamma \in \mathbb{R}_+^k$ and 0 elsewhere. In this case the posterior for β is a multivariate Student t density truncated to the positive orthant for γ. It may be the case that prior information is available on the other elements of β, say on the basis of economic theory. However, for the present discussion such extensions will not be pursued and a restricted uniform prior on β will be maintained.

Often one is interested in conducting conditional predictions of the total load for households with certain characteristics and certain appliance holdings. The CDA sampling model is extended to incorporate, say, N_f observations to be forecasted, denoted by y_f, conditional on an $N_f \times m$ design matrix X_f.

Equations (3.1) and (3.2) are modified to yield

(3.8) $y_s^e = X_s^e \beta + u_s^e$

and

(3.9) $p(u_s^e|\delta_s) = |V_s^e|^{-1/2} \delta_s^{T/2} g_s^e(\delta_s u_s^{e'}V_s^{e-1}u_s^e)$

with $T = N + N_f$, $y_s^{e\prime} = (y_s' \ y_f')$, $X_s^{e\prime} = (X_s' \ X_f')$, $u_s^{e\prime} = (u_s' \ u_f')$ and the $T \times T$ PDS matrix V_s^e is defined as

$$V_s^e = \begin{bmatrix} V_s & V_{sf} \\ V_{fs} & V_f \end{bmatrix}.$$

The function $g_s^e(.)$ should again satisfy the appropriate integrability condition and is linked to $g_s(.)$ in a way which does not depend on the parameters (β, δ_s); see for example Kelker (1970) or Dickey and Chen (1985).

Integrating out δ_s with the prior (3.7) and conditioning on β, y_s and X_s leads to

$$(3.10) \quad p(y_f | \beta, y_s, X_s^e) = f_s^{N_f} \left(y_f | N, \quad V_{fs} V_s^{-1} y_s + (X_f - V_{fs} V_s^{-1} X_s)\beta, \right.$$
$$\left. \frac{N}{(y_s - X_s\beta)' V_s^{-1} (y_s - X_s\beta)} V_{f.s}^{-1} \right)$$

with $V_{f.s} = V_f - V_{fs} V_s^{-1} V_{sf}$. We subsequently weigh (3.10) by the posterior density of β given all available sample information in order to obtain the post-sample predictive density:

$$(3.11) \quad p(y_f | y_s, X_s^e) = \int p(y_f | \beta, y_s, X_s^e) \, p(\beta | y_s, X_s) \, d\beta$$

A more detailed analysis is provided in Osiewalski and Steel (1992).

Specifying a noninformative prior for β implies that the predictive density given in (3.11) is a multivariate Student t while it becomes a Student t conditional on γ when γ is restricted to be nonnegative. The more general case can be examined by Monte Carlo integration, drawing directly from the posterior of β (i.e. direct sampling) and evaluating (3.10) at each value drawn for β. The latter is required for (marginal) density plots and quantiles of y_f given X_f and the observed sample, and could be useful in practice for moments of order three and higher. However, the first two

11

moments of y_f can easily be computed analytically from only the first two posterior moments of β if they exist, say $E(\beta|y_s, X_s)$ and $Var(\beta|y_s, X_s)$, using for $N > 1$

$$(3.12) \qquad E(y_f|y_s, X_s^e) = V_{fs}V_s^{-1}y_s + (X_f - V_{fs}V_s^{-1}X_s)\, E(\beta|y_s, X_s),$$

and if $N > 2$

$$(3.13) \qquad Var(y_f|y_s, X_s^e) = \frac{1}{N-2}\Bigg((N-m)s_s^2 + tr(X_s'V_s^{-1}X_s)\, Var(\beta|y_s, X_s) +$$

$$[\, E(\beta|y_s, X_s) - \hat{\beta}_s\,]'\, X_s'V_s^{-1}X_s\, [\, E(\beta|y_s, X_s) - \hat{\beta}_s\,]\Bigg) V_{f.s} +$$

$$(X_f - V_{fs}V_s^{-1}X_s)\, Var(\beta|y_s, X_s)\, (X_f - V_{fs}V_s^{-1}X_s)'.$$

Appendix I briefly indicates how (3.13) was obtained.

The 2-stage independent model
The most attractive feature of the CP model is that analytical results for posterior and predictive analyses are available. However this comes at the expense of the potentially restrictive assumption of assuming a common precision parameter for both the CDA and direct metering data. The 2-stage independent (2SI) model relaxes this assumption.

Basically, the idea is to conduct a sequential Bayesian analysis in which both types of sample information will in turn update our beliefs about the parameter of interest β. Precision parameters are allowed to be different for each type of sample information. This generalization comes at the expense of having to assume the independence of u_c and u_d. Correlation between these disturbances was permitted in the CP model albeit within the constraint of a known V_s matrix. Nevertheless the assumption of independence will be relaxed in our third model that utilizes a seemingly unrelated regression equation (SURE) framework to model the potential dependence.

In treating the CDA and direct metering data separately it needs to be

recognized that individually these data may not relate directly to all regression coefficients. Zero columns occur in (2.2) if say one or more appliances are fully metered while they occur in (2.3) when direct metering information is not available for the full set of k appliances under study. Thus, full-rank versions of (2.2) and (2.3) are written as follows:

(3.14) $y_c = X_c^* \beta_c^* + u_c$

and

(3.15) $y_d = X_d^* \beta_d^* + u_d$

where for $i = c, d$; $X_i^* = (Z_i^* \ D_i^*)$ is the $N_i \times (k_i^* + p_i^*)$ design matrix of full column rank obtained by taking the nonzero columns of X_i. Similarly β_i^* is a $(k_i^* + p_i^*)$ dimensional subvector of the m regression coefficients in β.

Again a multivariate elliptical assumption is made for the stochastic structure on both u_c and u_d so that

(3.16) $p(u_i | \delta_i) = |V_i|^{-1/2} \delta_i^{N_i/2} g_i(\delta_i u_i' V_i^{-1} u_i)$ $i = c, d$

where δ_i are positive scalar precision parameters, $g_i(.)$ is a known nonnegative function such that $r^{(N_i - 2)/2} g_i(r)$ is integrable in r over $\mathbb{R}_+$, V_i is a known $N_i \times N_i$ PDS matrix and u_c and u_d are assumed to be independent.

The prior density on the parameters $(\beta, \delta_c, \delta_d)$ is assumed to be

(3.17) $p(\beta, \delta_c, \delta_d) \propto \delta_c^{-1} \delta_d^{-1} I_+(\gamma).$

Although the ordering in this sequential Bayesian analysis is not well determined, given the independence between both sampling models, it seems natural to use the direct metering information in a first stage as in Hsiao et al. (1990). In any case, the results will not be influenced by the particular ordering chosen. Combining (3.17) with (3.15) and (3.16) then gives the following first-stage posterior for β:

(3.18) $p(\beta | y_d, X_d) = K_d^{-1} I_+(\gamma) f_s^{m_d} (\beta_d^* | N_d - m_d, \hat{\beta}_d^*, s_d^{*-2} X_d^{*\prime} V_d^{-1} X_d^*)$

13

where $m_d = k_d^* + p_d^*$ and where we use the hyperparameters

$$(3.19) \qquad \hat{\beta}_d^* = (X_d^{*'}V_d^{-1}X_d^*)^{-1}X_d^{*'}V_d^{-1}y_d$$

and

$$(3.20) \qquad s_d^{*2} = (y_d - X_d^*\hat{\beta}_d^*)' \, V_d^{-1}(y_d - X_d^*\hat{\beta}_d^*)/(N_d - m_d).$$

Note that the prior information on β is multiplied by the Student kernel in (3.18), which reflects the information contained in the directly metered sample about a subset of β. Conditionally upon this subset $\hat{\beta}_d^*$, the other elements of β are not affected by the sample information. The integrating constant K_d would be unity without the prior information on the sign of the γ_1, in which case the truncated Student in (3.18) would become a regular Student t density. Of course, (3.18) is only a proper posterior density for β_d^*, whereas the other elements of β retain their improper priors.

The beliefs about β summarized in (3.18) are posterior to the first stage, and it shall now be used as prior input for the second stage, namely the CDA analysis. As δ_c has not yet been a parameter of the sampling process, its prior independence of β is retained, and the second-stage prior is given by:

$$(3.21) \qquad p(\beta, \, \delta_c | y_d, \, X_d) \propto \delta_c^{-1} \, p(\beta | y_d, X_d)$$
$$\propto \delta_c^{-1} \, I_+(\gamma) \, f_s^{m_d}(\beta_d^* | \, N_d - m_d, \, \hat{\beta}_d^*, \, s_d^{*-2}X_d^{*'}V_d^{-1}X_d^* \,).$$

CDA data will then update the prior in (3.21) into a second-stage posterior of β, given by

$$(3.22) \qquad p(\beta | y_s, \, X_s) = K_c^{-1} \, I_+(\gamma) \, f_s^{m_c}(\beta_c^* | N_c - m_c, \, \hat{\beta}_c^*, \, s_c^{*-2}X_c^{*'}V_c^{-1}X_c^* \,)\times$$
$$f_s^{m_d} (\beta_d^* | N_d - m_d, \, \hat{\beta}_d^*, \, s_d^{*-2}X_d^{*'}V_d^{-1}X_d^* \,)$$

where m_c, $\hat{\beta}_c^*$ and s_c^{*2} are defined as in (3.18) - (3.20) with subscripts d replaced by c. The information from prior, first-stage sample and second-stage sample is clearly distinguishable in (3.22), which is a 2-0

14

poly-t density in common elements [see e.g. Drèze (1977)] truncated to the positive orthant for γ. It is assumed that the union of β_c^* and β_d^* is the entire β vector, since other elements could just be deleted from β in both sampling models. The intersection of both β^* vectors is typically non-empty, implying that certain appliances are directly metered for only part of the households in the CDA data set (i.e. partial direct metering) or that y_c and y_d come from different (but related) samples with overlapping appliance holdings. Those elements of β for which no direct metering information is gathered will have a truncated Student t posterior density, but *conditional* upon the other elements. Thus, the marginal posterior densities of *all* elements of β will be influenced by the direct metering information, even if the latter only applies to one or a few appliances.

The Student kernel in the first-stage posterior (3.18) is well-defined if

$$N_d > m_d,$$
$$(y_d - X_d^* \hat{\beta}_d^*)' \, V_d^{-1} (y_d - X_d^* \hat{\beta}_d^*) > 0, \text{ and}$$
$$X_d^{*'} V_d^{-1} X_d^* \text{ is PDS.}$$

However, formally these conditions are not required if our interest focuses on the second-stage posterior in (3.22). Drawings from the latter will be conducted as in Bauwens and Richard (1982) and truncation will be performed by a simple rejection approach. Apart from this truncation, (3.22) is a well-defined 2-0 poly-t density on the m-dimensional vector β if for i = c,d:

$$N > m$$
$$(y_i - X_i^* \hat{\beta}_i^*)' \, V_i^{-1} (y_i - X_i^* \hat{\beta}_i^*) > 0,$$
$$X_i' V_i^{-1} X_i \text{ is positive semi-definite symmetric (PSDS), and}$$
$$X_c' V_c^{-1} X_c + X_d' V_d^{-1} X_d \text{ is PDS,}$$

where the latter condition implies that every element of β should appear in either β_c^* or β_d^*. Posterior moments of β will exist at least up to order $\rho <$ N - m.

If our interest is in predicting the load of one particular directly metered appliance, the direct metering model in (3.15) and (3.16) should be used. If

on the other hand conditional predictions of the total load for households with certain characteristics and certain appliance holdings is required then it seems more relevant to use the CDA model in (3.14) and (3.16). Thus the precision associated with that type of aggregate measurement is used and it is implicitly assumed that $\delta_d V_d^{-1}$, the spread of the direct metering process, is dominated by $\delta_c V_c^{-1}$, its counterpart for the CDA process.

The CDA sampling model is extended to incorporate the N_f observations to be forecasted to yield

(3.23) $y_c^e = X_c^e \beta + u_c^e$

and

(3.24) $p(u_c^e | \delta_c) = |V_c^e|^{-1/2} \delta_c^{N_c^e/2} g_c^e(\delta_c u_c^{e\prime} V_c^{e-1} u_c^e)$

with $N_c^e = N_c + N_f$, $y_c^{e\prime} = (y_c{}^\prime \ y_c^{f\prime})$, $X_c^{e\prime} = (X_c{}^\prime \ X_c^{f\prime})$, $u_c^{e\prime} = (u_c{}^\prime \ u_c^{f\prime})$ and the $N_c^e \times N_c^e$ PDS matrix V_c^e is defined as

$$V_c^e = \begin{bmatrix} V_c & V_{cf} \\ V_{fc} & V_f \end{bmatrix}.$$

Conditions on $g_c^e(.)$ are as they were for $g_s^e(.)$.

As before δ_c is integrated out and we condition on β, y_s, X_s and X_c^f to yield

(3.25) $p(y_c^f | \beta, X_c^f, y_s, X_s) = p(y_c^f | \beta, X_c^e, y_c)$

$$= f_s^{N_f} \left(y_c^f | N_c, \ V_{fc} V_c^{-1} y_c + (X_c^f - V_{fc} V_c^{-1} X_c)\beta, \ \frac{N_c}{(y_c - X_c \beta)^\prime V_c^{-1}(y_c - X_c \beta)} V_{f.c}^{-1} \right)$$

with $V_{f.c} = V_f - V_{fc} V_c^{-1} V_{cf}$. We weigh (3.25) by the posterior density of β given all available sample information, i.e. the second-stage posterior given by (3.22), in order to obtain the post-sample predictive density:

$$(3.26) \qquad p(y_c^f | X_c^f, \ y_s, \ X_s) = \int p(y_c^f | \beta, \ X_c^f, \ y_s, \ X_s) \ p(\beta | y, \ X) \ d\beta$$

Clearly, those elements of β that are not truncated and only appear in the CDA sample can be integrated out analytically in (3.26), given the rest of β. Only if there is no direct metering and we do not impose the non-negativity constraints on γ, do we obtain a fully analytical solution to (3.26) in a Student t form. For the more general case Monte Carlo integration is required in order to determine (marginal) density plots and quantiles of y_c^f given X_c^f and the observed sample, and could be useful in practice for moments of order three and higher.

However, as with the CP model the first two moments of y_c^f can easily be computed analytically from only the first two posterior moments of β if they exist, using for $N_c > 1$

$$(3.27) \qquad E(y_c^f | X_c^f, \ y_s, \ X_s) = V_{fc} V_c^{-1} y_c + (\ X_c^f - V_{fc} V_c^{-1} X_c) \ E(\beta | y_s, \ X_s),$$

and if $N_c > 2$

$$
\begin{aligned}
(3.28) \qquad \mathrm{Var}(y_c^f | X_c^f, \ y_s, \ X_s) = \ &\frac{1}{N_c - 2} \Bigg((N_c - m_c) s_c^{*2} + \\
&\mathrm{tr}(\ X_c^{*\prime} V_c^{-1} X_c^{*}) \ \mathrm{Var} \ (\beta_c^{*} | y_s, \ X_s) + \\
&[E(\beta_c^{*} | y_s, \ X_s) - \hat{\beta}_c^{*} \ ]' \ X_c^{*\prime} V_c^{-1} X_c^{*} \ [\ E(\beta_c^{*} | y_s, \ X_s) - \hat{\beta}_c^{*} \ ] \Bigg) V_{f \cdot c} + \\
&(X_c^f - V_{fc} V_c^{-1} X_c) \ \mathrm{Var}(\beta | y_s, \ X_s) \ (X_c^f - V_{fc} V_c^{-1} X_c)'.
\end{aligned}
$$

Even for the homoskedastic sampling model obtained by taking V_c^e to be an identity matrix in (3.24), the uncertainty about β will introduce heteroskedasticity into the predictive density through the last term in (3.28). In this special case the error vector u_c^e is said to have a multivariate spherical distribution; see Kelker (1970).

The 2-stage SURE model

In the 2SI model, disturbances associated with the direct metering data are assumed to be independent of those disturbances associated with the CDA data. This is potentially restrictive especially in the situation where some households appear in both the CDA and direct metering samples. Write u_c in terms of two subvectors: u_{cD} referring to directly metered households in the CDA sample and u_{cn} referring to the remaining households in the CDA sample. The 2-stage SURE (2SS) model allows for correlation between u_{cD} and u_d but requires the independence of u_{cn} with u_{cD} and u_d. For the directly metered households the CDA and direct metering models will form a two-equation SURE model. Because of the structure of our models the correlation is between observations on the same household and hence represents the presence of common household unobservables.

As in the 2SI model, the approach is to conduct a sequential Bayesian analysis. Here though, the first stage analysis is a CDA model for the households for which there is no direct metering data. This then provides the prior for a SURE second stage involving the directly metered households for which there are both CDA and direct metering observations.

The partitioning of the data between the first and second stages of the 2SS model suggests an alternative form of the 2SI model. Rather than using the direct metering data as the first stage followed by the CDA data, the alternative 2SI model partitions the data between stages according to whether the household was directly metered or not. Just as in the 2SS, the first stage analysis is a CDA model for the households with no direct metering data. The second stage then utilizes the CDA and direct metering observations available for the directly metered households, but not in a SURE model. Because the 2SI model resulting from the 2SS partitioning requires a common precision parameter for the direct metering and CDA data of the metered households, it is unlikely to be preferred in applications to either our 2SS or original 2SI models.

Using now familiar notation the first stage model is represented by

$$(3.29) \qquad y_{cn} = X_{cn}^* \beta_{cn}^* + u_{cn}$$

and u_{cn} is assumed to be multivariate elliptical with location zero and scale matrix given by $\delta_{cn}^{-1} V_{cn}$. X_{cn}^{*} is a $N_{cn} \times m_{cn}$ design matrix of full column rank and β_{cn}^{*} is a conformable subvector of β.

With the prior density assumed to be

$$(3.30) \qquad p(\beta, \delta_{cn}) \propto \delta_{cn}^{-1} I_{+}(\gamma),$$

the first-stage posterior for β has exactly the same form as the corresponding first-stage posterior for the 2SI model. As such the expressions given in (3.18) – (3.20) are appropriate after replacing the subscript "d" by "cn".

For the second stage

$$(3.31) \qquad y_{cD} = X_{cD}\beta + u_{cD} = X_{cD}^{*}\beta_{cD}^{*} + u_{cD}$$
and
$$(3.32) \qquad y_{d} = X_{d}\beta + u_{d} = X_{d}^{*}\beta_{d}^{*} + u_{d}.$$

It is not possible to retain the generality of elliptical distributions in this SURE framework. Consequently $U_{D} = (u_{cD} \ u_{d})$ is assumed to be distributed as a multivariate normal with mean zero and covariance matrix $\Sigma \otimes V_{D}$. Σ is a 2×2 PDS covariance matrix that incorporates the potential correlation between u_{cD} and u_{d} while V_{D} is a known $N_{d} \times N_{d}$ PDS covariance matrix that is common for u_{cD} and u_{d}.

A Jeffrey's prior for Σ is assumed which in our 2-equation SURE is given by

$$(3.33) \qquad p(\Sigma) \propto |\Sigma|^{-3/2}$$

Let $Y_{D} = (y_{cD} \ y_{d})$ and $X_{D} = (X_{cD} \ X_{d})$ then the posterior for Σ is given by

$$(3.34) \qquad p(\Sigma | \beta, Y_{D}, X_{D}) = f_{IW}(\Sigma | U_{D}'V_{D}^{-1}U_{D}, N_{d})$$

where $f_{IW}(.)$ is an inverted Wishart density and

$$U_D = Y_D - X_D(\beta \; \beta) = Y_D - (X^*_{cD} \; X^*_d) \begin{pmatrix} \beta^*_{cD} & 0 \\ 0 & \beta^*_d \end{pmatrix}.$$

The marginal posterior for β can be written as

$$(3.35) \qquad p(\beta | y_s, X_s) \propto I_+(\gamma) \, f_s^{m_{cn}}(\beta^*_{cn} | N_{cn} - m_{cn}, \, \hat{\beta}^*_{cn}, \, s^{*-2}_{cn} X^{*\prime}_{cn} V^{-1}_{cn} X^*_{cn}) \times$$
$$|U_D' V^{-1}_D U_D|^{-N_d/2}.$$

The last factor in (3.35) can be written as a Student density in β^*_{cD} given β^*_d, times a 2-1 poly-t density in β^*_d; see Drèze and Morales (1976). Thus, (3.35) will generally be the product of a truncated conditional 2-0 poly-t density and a truncated marginal 3-1 poly t density. In the case of complete direct metering we have a 2-0 poly-t density times a 2-1 poly-t density both truncated. For each of these cases the properties of (3.35) will need to be investigated by means of Monte Carlo integration.

Prediction involves extending the first stage model representing households that were not directly metered. Conceptually it is identical to the procedure outlined for prediction in the 2SI model. With N_f observations to be forecasted, equations (3.23) to (3.26) are appropriate after systematically replacing subscript 'c' by 'cn'. Because of the form of the posterior for β given in (3.35), evaluation of the 2SS post-sample predictive density will need to be done numerically.

IV An application to Australian electrical appliance loads

Data and initial results

Data for this study were compiled as part of the Domestic End-Use Study (DES) conducted over a period of 15 months in 1986-87 in the state of New South Wales (NSW), Australia, under the auspices of an industry committee representing all the relevant authorities in NSW. The same data set was used by Fiebig et al. (1991) to study variations in household appliance loads over the day. Here we concentrate on monthly electricity consumption and in particular on the differences between end-use consumptions on weekdays versus weekends. For the purposes of illustration our attention is

confined to July, the winter month of peak demand.

Table 1 provides a brief description and some relevant summary statistics of the variables used in the analysis. The sample consists of 174 households located in the Sydney area. Direct metering observations are available for 2 appliances: HWPK, 21 observations; and HWOP, 87 observations. In the case of HWOP all households within the sample that possess this particular end-use have been metered.

Notice that the workday consumption has lower mean and variance than weekend consumption. Our analysis is designed to make similar comparisons between the end-use consumptions rather than their totals and to identify the source of the differences in total consumption levels.

For three of the end-uses the UEC's of (2.1) are assumed to vary with household characteristics. In particular the UEC's of HWPK and HWOP are assumed to be linearly related to PEOPLE while that of HEAT is assumed to be a linear function of SIZE. This implies inclusion of the following interaction variables into our CDA specification: SIZE*HEAT, PEOPLE*HWPK and PEOPLE*HWOP. PEOPLE, SIZE, INCOME, INCMSG and MJROTH are also included in the specification: they can be thought of as interactions with the INTERCEPT which is interpreted as the consumption for the average household of all common end-uses such as lighting, washing machines, etc. This interpretation derives from the measurement of all demographic variables in terms of deviations from the mean. The UEC's of all other appliances are taken to be constant over households.

Complete specifications of the regression models are provided in Table 2 where initial OLS regression estimates are given. Incorporating the available metering information requires separating the direct metering consumptions from total consumptions. The first set of results represent the estimated version of equation (2.2) which has as its dependent variable total consumption net of direct metering. Because HWOP is totally metered this dummy and the associated interaction do not appear in this equation. The second set of results are for the estimated version of equation (2.3) where direct metering consumptions are regressed on the two HWPK and HWOP dummies and the associated interactions.

21

These models and their associated estimates are provided for the purposes of illustration and to highlight a number of deficiencies that need to be addressed. Available direct metering information is not fully utilized as the commonality of coefficients in (2.2) and (2.3) has been ignored. There is nothing here to ensure the non-negativity of the estimated end-use consumptions (see FREEZ and POOLPUMP) nor will conventional confidence intervals for positive estimated end-uses necessarily have positive lower limits (see FRIGAUT, DRYER and MWAVE, and to a lesser extent DSH). Our proposed Bayesian models provide a framework for resolving these problems. Furthermore, heteroskedasticity in the disturbances will be accounted for within this framework.

Bayesian results

The choice between the alternative Bayesian models will depend primarily on how appropriate their different stochastic structures are for the particular problem being investigated. Disturbances associated with the CDA data for non-metered households are expected to be independent of all other disturbances. Moreover, using directly metered households only, simple correlations are small between the OLS residuals from the CDA and direct metering regressions; -0.046 for workday data and -0.006 for weekend data. This suggests that an assumption of independent disturbances may be satisfactory for these data. Consequently the extra computational effort required for the 2SS model does not seem warranted and our attention will be confined to the CP and 2SI models.

Recall that with the CP model any heteroskedasticity, which does seem to be a problem here, needs to be modelled in terms of a known V_s matrix. Alternatively we must be willing to condition on any unknown parameters in V_s. A convenient and reasonable assumption to make here is that heteroskedasticity is proportional to the household's appliance holdings. More appliances imply greater consumption and hence greater variability in electricity demand. Thus V_s is specified as a diagonal matrix with diagonal elements proportional to the number of appliances held by the particular household. This parameterization of heteroskedasticity is retained for the 2SI model. However recall that the 2SI specification involves relaxing the potentially restrictive assumption made in the CP model that there is a

22

common precision parameter for both the CDA and direct metering data.

As mentioned previously an attractive feature of the CP model is the availability of analytical results for posterior and predictive analyses. For applied work this feature translates into computational convenience. In particular the CP model is readily accessible to the practitioner because the popular econometric software package SHAZAM, (see White (1978)), can be used for calculations. Unfortunately the capabilities of SHAZAM do not extend to the estimation of the 2SI model. Consequently, the reported CP results were generated by SHAZAM while the 2SI results were generated by a computer program for Monte Carlo integration described in the computational appendix.

Table 3 reports posterior means and standard deviations for the CP and 2SI models. Results that do not incorporate the non-negativity constraint were also produced but for most coefficients the effect of truncation is minor. As an example consider 2SI estimation of the COOK parameter for workdays. Marginal posterior densities with and without truncation provided in Figure 1 are unimodal, essentially symmetric and almost the same. One of the few cases where truncation is influential is illustrated in Figure 2 which provides the 2SI marginal posteriors with and without truncation for the MWAVE end-use on workdays. Even though the posterior mean without truncation is positive, the effect of truncation is dramatic producing a highly skewed marginal posterior density. A complete set of CP and 2SI results without truncation are available on request.

Differences between CP and 2SI estimates of individual end-use demands are typically small. This is not surprising since there is little conflict of information for HWPK and PEOPLE*HWPK coefficients in the CDA and direct metering samples. This can be seen by comparing the corresponding OLS estimates given in Table 2.

Taking the 2SI results as the basis of further discussion we first note that differences between the end-use demands of workdays and weekends are rarely large in magnitude. The most prominent differences are found in the HEAT and INTERCEPT demands. (Recall that the estimate associated with the INTERCEPT can be interpreted as electricity consumption of the average

23

household attributable to appliances with high penetrations such as lighting.) For non-discretionary appliances such as FREEZ, HWPK and HWOP this similarity in demands for the different daytypes is expected but for the other appliances where there is more scope for discretionary use it represents an interesting result.

Although differences in the two sets of end-use demand estimates are not overly large the pattern of differences is striking. With the exception of HEAT and POOLPUMP, the latter being a minor exception, all workday end-use demands are less than the corresponding weekend demands. Also the variabilities of the workday estimates are uniformly lower. This replicates the previously identified relationship between the total demands. These features of the results are presumably indicative of differences in the daytypes such as the amount of time spent at home.

A comparison of predictive densities for particular household types further illustrates the differences in daytypes. As an example consider a "typical" electric household. Possession of electric water heating (either peak or off-peak) is taken to define electric households. Then amongst this sub-group of households an appliance portfolio was chosen to include the appliances with the four highest penetrations. Consequently our chosen household is characterized by HWOP=1, HEAT=1, DRYER=1, and COOK=1; in addition we set INTERCEPT=1, and PEOPLE, SIZE and SIZE*HEAT at the average values for those households with HWOP=1. All other demographic variables are set to the sample mean values of zero.

Figure 3 provides the predictive densities for the two daytypes calculated with truncation. Both the workday and weekend day predictive densities are symmetric but the former has a smaller expected value (39.65 compared with 41.15) and a smaller standard deviation (11.94 compared with 13.20). What is also clear from both densities is that there is considerable dispersion in the predicted consumption of our typical electric household.

Because our typical electric household includes end-use estimates that individually were not overly affected by whether estimates were truncated or not, one would not expect substantial differences between the densities in Figure 3 and those calculated without truncation. A comparison of these

predictive densities confirmed this but it is noteworthy that the truncated densities were shifted to the left and they exhibited less dispersion.

V Conclusion

Estimates that fail to satisfy strongly held a priori restrictions are difficult to justify. Negative estimates of electricity end-use or appliance demand are a good example of this. Estimating these consumptions by direct metering methods solves the problem but at a large cost. Moreover, such methods ignore readily available information from billing records and appliance-holding surveys. We have demonstrated how these two aspects of estimating end-use demand can be readily integrated into a Bayesian framework. Our discussion has also targetted two potential extensions to the Bayesian models that have been developed. A richer class of heteroskedasticity or correlation between disturbances would require the V matrix to depend on unknown parameters. This extension requires an added computational burden but little in the way of analytical development. The second extension involves the relaxation of the assumption that appliance ownership is exogenously determined. This represents a more challenging task that must be left for future work.

Appendix I: Derivation of (3.13)

In order to calculate the marginal predictive variance in (3.13) we have used the well-known rule that:

$$\text{Var}(y_f|y_s, X_s^e) = E_\beta \left[\text{Var}(y_f|\beta, y_s, X_s^e) \right] + \text{Var}_\beta \left[E(y_f|\beta, y_s, X_s^e) \right],$$

in which the first term involves the mean of the quadratic form

$$E_\beta \left[(y_s - X_s\beta)'V_s^{-1}(y_s - X_s\beta) \right] = (N - m)s_s^2 + \text{tr}(X_s'V_s^{-1}X_s) \ \text{Var}(\beta|y_s, X_s)$$

$$+ [E(\beta|y_s, X_s) - \hat{\beta}_s]' \ X_s'V_s^{-1}X_s \ [E(\beta|y_s, X_s) - \hat{\beta}_s]$$

where $\hat{\beta}_s$ and s_s^2 are defined as in (3.5) and (3.6).

Appendix II: Computational Appendix

The "MCI" (Monte Carlo Integration) program described in Bauwens and Richard (1982, Section 5) has been used for computing the Bayesian results. The main ingredient of the program is an algorithm (given in their paper) for generating random drawings of a 2-0 poly-t density (for the 2SI model). Each drawing of β that falls in the region where $I_+(\gamma) = 1$ is a drawing of the truncated density, see (3.22), and is used for computing the posterior expectation and covariance matrix of β (by simple averaging over a predetermined number M of drawings). Given these results and the required inputs, the predictive mean and variance can be computed by a few matrix operations, see (3.27) and (3.28). The generated values $\beta_{(l)}$ are also used to compute the predictive density (3.26). For a given value of y_c^f, the right-hand side of (3.26) is approximated by:

$$M^{-1} \sum_1 p(y_c^f | \beta_{(l)}, X_c^f, y_s, X_s) \ ,$$

i.e. the simple average of the conditional densities evaluated at y_c^f. To plot the predictive density, this is done for a grid of y_c^f values (33 for the reported plots), at the same time that the posterior moments and marginal densities of β are computed.

The reported results were computed on a PC equipped with a 80486DX processor running at 33 Mhz. The program is written in FORTRAN 77 using the compiler of MICROSOFT (version 4.10). Basic results (such as reported in Table 3 for one day type, 2SI model) can be obtained with high accuracy by requiring

26

10,000 drawings from the truncated posterior density. In our application, the rejection rate was about 16 to 1, i.e. 160,000 drawings from the untruncated density were generated, out of which 150,000 had to be discarded. The computing time for this is almost one hour. In order to get accurate graphs of marginal densities of elements of β (based on 33 points which can be used for interpolating the curves), we used 150,000 drawings; the computing time was therefore about 15 hours (one night). For minimizing the computing time, it is essential to test for the positivity restrictions on the γ coefficients in the best order; this can be determined by computing the untruncated posterior moments in a preliminary run. In order to appreciate the computing time of our application, recall that the density to be simulated is of dimension 19, hence 150,000 accepted drawings require 48 million uniform random numbers!

References

Aigner, D.J., Sorooshian, C. and Kerwin, P. (1984). "Conditional demand analysis for estimating residential end-use load profiles", The Energy Journal, 5, 81-97.

Bartels, R. and Fiebig, D.G. (1990). "Integrating direct metering and conditional demand analysis for estimating end-use loads", The Energy Journal, 11, 79-97.

Bauwens, L. and Richard, J.-F. (1982). "A poly-t random variable generator, with application to Monte Carlo integration", CORE Discussion Paper no. 8214.

Bauwens, L. and Lubrano, M. (1991). "Bayesian diagnostics for heterogeneity", Annales d'Economie et de Statistique, 20/21, 17-40.

Caves, D.W., Herriges, J.A., Train K.E. and Windle, R.J. (1987). "A Bayesian approach to combining conditional demand and engineering models of electricity usage", Review of Economics and Statistics, 69, 438-448.

Dickey, P. and Chen, C.-H. (1985). "Direct subjective-probability modelling using ellipsoidal distributions", in J.M. Bernado, M.H. DeGroot, D.V. Lindley and A.F.M. Smith, eds., Bayesian Statistics 2 (North-Holland, Amterdam), 57-182.

Drèze, J.H. (1977). "Bayesian regression analysis using poly-t densities", Journal of Econometrics, 6, 329-354.

Drèze, J.H. and Morales, J.A. (1976). "Bayesian full information analysis of simultaneous equations", Journal of the American Statistical Association, 71, 919-923.

Engle, R.F., Hendry, D.F. and Richard, J.-F. (1983). "Exogeneity", Econometrica, 51, 277-304.

Fiebig, D.G., Bartels, R. and Aigner, D.J. (1991). "A random coefficient approach to the estimation of residential end-use load profiles", Journal of Econometrics, 50, 297-328.

Geweke, J. (1986). "Exact inference in the inequality constrained normal linear regression model", Journal of Applied Econometrics, 1, 127-141.

Hsiao, C., Mountain, D.C. and Ho, K.F. (1990). "A Bayesian integration of end-use metering and conditional demand analysis", Paper presented at the 6th World Congress of the Econometric Society in Barcelona.

Johnson, M.E. (1987). Multivariate Statistical Simulation, New York: Wiley.

Kelker, D. (1970). "Distribution theory of spherical distributions and a location-scale generalization", Sankhya A, 32, 419-430.

Koenker, R. (1981). "A note on studentizing a test for heteroskedasticity", Journal of Econometrics, 17, 107-112.

Osiewalski, J. and Steel, M. (1992). "Robust Bayesian inference in elliptical regression models", forthcoming Journal of Econometrics.

Sebold, F.D. and Parris, K.M. (1989). "Residential end-use energy consumption: A survey of conditional demand estimates", Report CU-6487, Electric Power Research Institute, Palo Alto, CA.

White, K.J. (1978). "A general computer program for econometric methods - SHAZAM", Econometrica, 46, 239-240.

Table 1: Variable descriptions and summary statistics

Variables	Max	Min	Mean	Std Dev
Appliance indicators				
FREEZ = separate freezer	1	0	.38	.49
FRIGAUT = automatic defrost fridge	1	0	.59	.49
COOK = electric oven or hotplates	1	0	.72	.45
DSH = dishwasher	1	0	.33	.47
DRYER = clothes dryer	1	0	.56	.50
HEAT = electric main or secondary heating	1	0	.84	.36
HWPK = main tariff water heater	1	0	.18	.39
HWOP = offpeak tariff water heater	1	0	.50	.50
POOLPUMP = pool pump	1	0	.11	.31
MWAVE = microwave oven	1	0	.39	.49
MJROTH = other major appliance, e.g. spa	1	0	.06	.23
Demographic variables				
SIZE = area of home	37.5	4	15.64	8.14
PEOPLE = number of household members	8	1	2.99	1.63
INCOME = annual before tax income ($10,000)	9.0	.45	3.38	2.43
INCMSG = dummy variable for missing income	1	0	.14	.35
Consumption variables[*]				
Net consumption excluding metering:				
Workday	85.7	1.5	23.0	16.9
Weekend	94.0	1.4	24.5	17.9
Metered consumption of HWOP:				
Workday	37.7	4.7	14.9	6.9
Weekend	43.5	5.2	15.3	7.7
Metered consumption of HWPK:				
Weekday	22.7	2.9	12.9	5.5
Weekend	29.6	1.8	13.2	6.0

[*] Of the 174 households in the sample, 21 have metered HWPK consumption and 87 have metered HWOP consumption. Summary statistics of the metered consumptions were calculated over the metered subsamples while those of INCOME were calculated over households with income information.

Table 2: OLS regression estimates[*]

Explanatory variables	CDA using net consumption		Direct metering	
	Workday	Weekend	Workday	Weekend
INTERCEPT	4.93	5.85		
	(2.98)	(3.36)		
FREEZ	-1.92	-1.59		
	(2.20)	(2.48)		
FRIGAUT	2.11	2.01		
	(2.08)	(2.34)		
COOK	6.43	7.24		
	(2.20)	(2.48)		
DSH	4.11	5.40		
	(2.56)	(2.88)		
DRYER	2.97	4.06		
	(2.40)	(2.70)		
HEAT	9.78	8.56		
	(2.69)	(3.03)		
HWPK	8.32	7.44	11.45	11.85
	(2.58)	(2.91)	(1.38)	(1.60)
HWOP			14.87	15.28
			(0.62)	(0.72)
POOLPUMP	-0.87	-2.40		
	(3.29)	(3.71)		
MWAVE	0.72	0.84		
	(2.09)	(2.35)		
PEOPLE	3.05	2.94		
	(0.71)	(0.80)		
SIZE	-0.40	-0.48		
	(0.37)	(0.41)		
INCOME	0.67	0.90		
	(0.48)	(0.54)		
INCMSG	1.83	4.59		
	(3.19)	(3.59)		
MJROTH	3.16	1.66		
	(4.36)	(4.92)		
PEOPLE*HWPK	2.10	2.29	1.81	1.74
	(1.68)	(1.90)	(0.72)	(0.84)
PEOPLE*HWOP			2.32	2.36
			(0.42)	(0.48)
SIZE*HEAT	1.06	1.00		
	(0.38)	(0.42)		
$\bar{R}^2$	0.51	0.44	0.25	0.20
Breusch-Pagan	9.54	16.17	0.60	1.78
Studentized Breusch-Pagan[**]	5.59	8.35	0.34	0.77

[*] Numbers in brackets are standard errors.

[**] The Breusch-Pagan and Studentized Breusch-Pagan (see Koenker (1981)) test for heteroskedasticity related to total appliance holdings. Both are asymptotically Chi-squared with one degree of freedom.

Table 3: Bayesian estimates[*]

Explanatory variables	Common Precision		2-stage Independent	
	Workday	Weekend	Workday	Weekend
INTERCEPT	5.08	6.12	4.97	5.96
	(2.30)	(2.59)	(2.18)	(2.41)
FREEZ	1.33	1.63	1.31	1.56
	(1.10)	(1.32)	(1.06)	(1.25)
FRIGAUT	2.05	2.48	1.98	2.30
	(1.40)	(1.63)	(1.34)	(1.52)
COOK	5.02	5.58	5.33	6.06
	(1.88)	(2.08)	(1.78)	(1.97)
DSH	3.34	4.09	3.18	3.90
	(2.15)	(2.52)	(1.98)	(2.30)
DRYER	2.22	2.78	2.20	2.74
	(1.51)	(1.83)	(1.48)	(1.75)
HEAT	9.09	7.87	9.37	8.20
	(2.38)	(2.67)	(2.23)	(2.46)
HWPK	10.71	10.89	10.57	10.64
	(1.08)	(1.24)	(1.18)	(1.35)
HWOP	14.87	15.29	14.87	15.28
	(0.58)	(0.65)	(0.62)	(0.72)
POOLPUMP	3.13	3.14	3.04	2.87
	(2.34)	(2.42)	(2.29)	(2.27)
MWAVE	1.97	2.32	1.99	2.28
	(1.40)	(1.65)	(1.40)	(1.58)
PEOPLE	2.49	2.44	2.47	2.42
	(0.65)	(0.75)	(0.61)	(0.68)
SIZE	-0.22	-0.26	-0.23	-0.27
	(0.32)	(0.36)	(0.30)	(0.33)
INCOME	0.48	0.70	0.48	0.70
	(0.47)	(0.54)	(0.46)	(0.50)
INCMSG	1.88	3.67	1.69	3.50
	(2.71)	(3.09)	(2.71)	(2.99)
MJROTH	1.43	-0.50	0.92	-0.98
	(4.84)	(5.50)	(4.42)	(4.90)
PEOPLE*HWPK	2.10	2.10	2.13	2.18
	(0.60)	(0.67)	(0.64)	(0.73)
PEOPLE*HWOP	2.32	2.36	2.32	2.36
	(0.38)	(0.43)	(0.42)	(0.49)
SIZE*HEAT	0.90	0.87	0.91	0.89
	(0.33)	(0.38)	(0.32)	(0.35)

[*] Numbers in brackets are standard deviations.
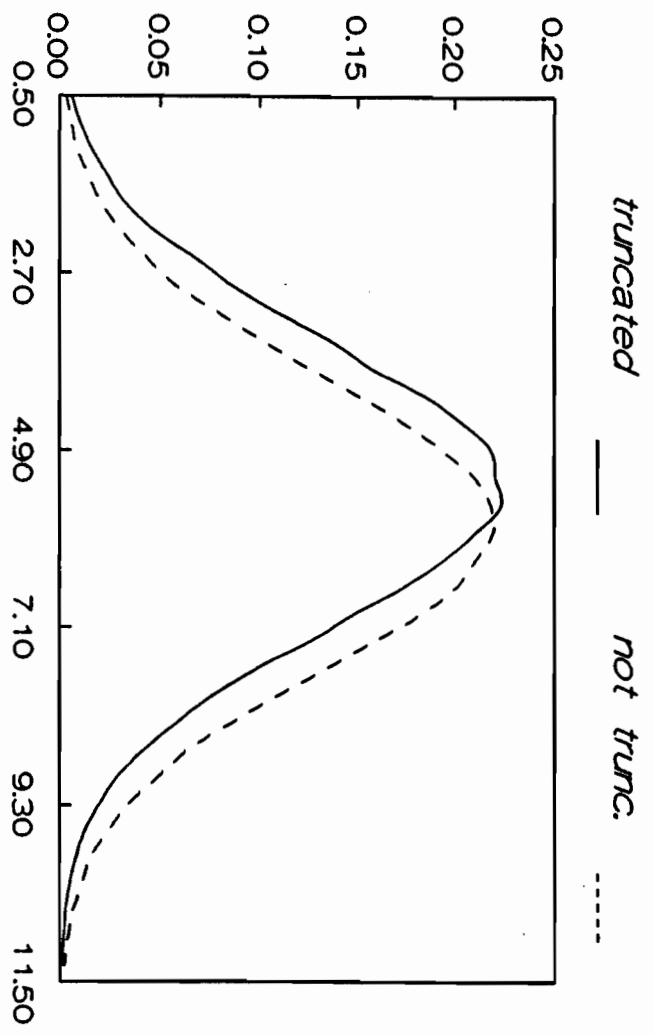
Fig. 1: Posterior density of COOK
workday data, 2SI

truncated ——

not trunc. ·····

Fig. 2: Posterior density of MWAVE
workday data, 2SI

truncated ———   not trunc. ·····

0.00

0.06

0.12

0.18
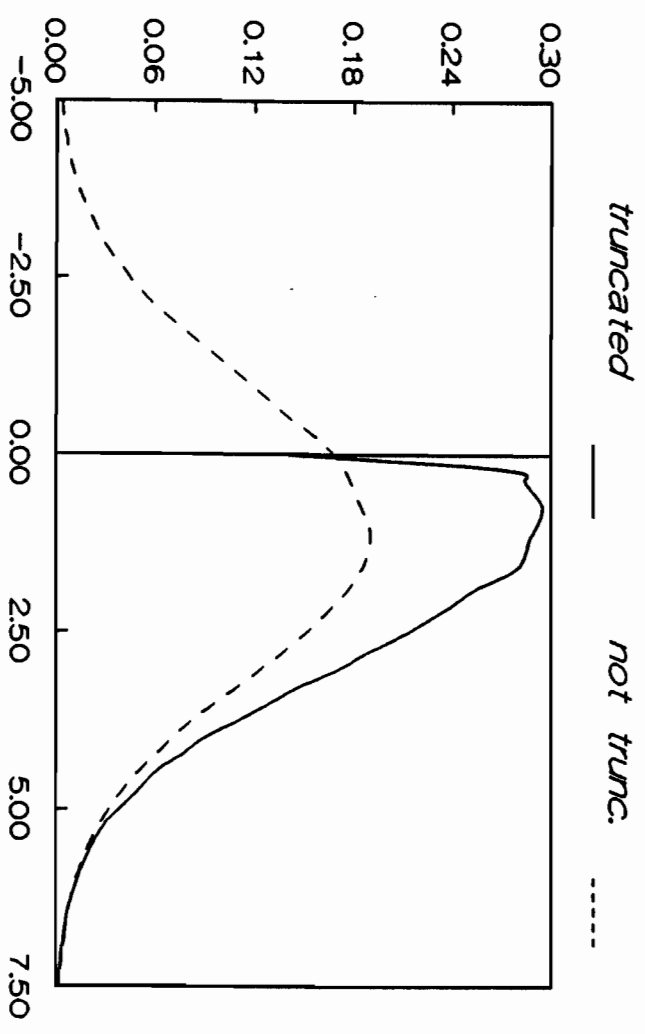
0.24

0.30

-5.00   -2.50   0.00   2.50   5.00   7.50

Fig. 3: Predictive density
typical household, truncated

workday ——   weekend ----