Integrating Direct Metering and Conditional Demand Analysis for Estimating End-Use Loads

Bartels, R.; Fiebig, D.

Publication date:
1990

entER
for
mic Research

Discussion paper

TILBURG UNIVERSITY

No. 9056

INTEGRATING DIRECT METERING AND CONDITIONAL DEMAND ANALYSIS FOR ESTIMATING END-USE LOADS

by Robert Bartels
and Denzil G. Fiebig

September 1990

ISSN 0924-7815

INTEGRATING DIRECT METERING AND CONDITIONAL DEMAND ANALYSIS FOR ESTIMATING END-USE LOADS*

Robert Bartels
University of Sydney and University of Bonn

Denzil G. Fiebig
University of Sydney and CentER at Tilburg University

October 1989
Revised August 1990

Abstract

Conditional demand analysis (CDA) is a statistical method for allocating the total household electricity load during a period into its constituent components, each associated with a particular electricity-using appliance or end-use. This is an indirect approach to the estimation of end-use demand and quite naturally it often generates imprecise estimates. One of the possible methods for improving these estimates involves the incorporation of data obtained by directly metering specific appliances. It is argued that an extremely natural approach to the use of this extra information follows directly from a reformulation of the standard CDA model into a random coefficient framework. Some new results on the possible efficiency gains from such an approach are developed. Illustrations based on an empirical study of households in the state of New South Wales, Australia are also provided.

* This paper developed from a project carried out in the state of New South Wales in Australia in conjunction with an Electricity Industry Working Group (IWG) comprising representatives from the Electricity Commission of N.S.W., the N.S.W. Department of Energy, the Local Government Energy Association and the County Councils of Sydney, Prospect, Southern Riverina, Illawarra and Shortland. The authors gratefully acknowledge the assistance of the IWG. We are especially indebted to the Commission and its officers Bob Lumsdaine and Mike Garben whose assistance was invaluable in the preparation of this paper. We also wish to thank Dr. Mark Steel of Tilburg University and Dr. Joachim Werner of Bonn University for helpful discussions, Kerrie Legge for providing excellent research assistance and an anonymous referee for some helpful suggestions. A preliminary version of the paper was presented at the Australasian Meetings of the Econometric Society in Armidale, July 1989.

1. INTRODUCTION

Conditional demand analysis (CDA) is a statistical method for allocating the total household electricity load during a period into its constituent components, each associated with a particular electricity-using appliance or end-use. The method, introduced by Parti and Parti (1980), assumes that the electricity load of a household is linked via a linearly additive regression model to a set of dummy variables representing the household's appliance ownership. The estimated coefficients of the dummy variables can be interpreted as the mean contribution of each end-use to the total load. End-use load profiles through the day can be obtained by applying CDA for each hour through the day, an approach pioneered by Aigner, Sorooshian and Kerwin (1984).

As revealed in a recent investigation by Huss (1985), accurate information about end-use loads is increasingly seen by electricity utilities as important for their generation planning, marketing and rate-making activities. However, CDA is an indirect approach to the estimation of end-use loads and the estimates it produces are not always precise or plausible. Negative loads, or technically implausibly large loads, are difficult to justify. The problem primarily arises because the ownership of appliances among households in the sample is generally not very heterogeneous.

To overcome this weakness of CDA it is natural to look for additional sources of information with which to supplement the analysis. Caves, Herriges, Train and Windle (1987) consider end-use profiles produced by engineering models based on thermodynamic principles, and propose a Bayesian approach for combining these profiles with CDA. Engineering models are only appropriate, however, in situations where individual behavior plays a minor role, for example, heating and cooling in extreme climates. Most appliance use depends on the life style; in temperate climates, even heating and cooling appliances are in many households only used when the occupants are at home.

An alternative way of obtaining additional information about end-use electricity consumption is to meter specific appliances directly. Ad hoc direct metering of specific appliance types is sometimes carried out by utilities to assist in developing marketing strategies. As technical advances reduce the cost of such metering, this option is now also becoming increasingly attractive as a way of supplementing the usual household load data collected for load research purposes.

Recently, methods of integrating the data gathered using direct metering on selected appliances in a sample of households in a load research study with CDA have been suggested. The basis for our discussion is the Fiebig, Bartels and Aigner (1988) approach that follows directly from a reformulation of the standard CDA model into a random coefficient framework. However, an alternative approach suggested by Aigner and Schönfeld (1988), Caves et al. (1987) and Hsiao, Mountain and Ho (1990) where the direct metering information is treated as stochastic prior information is shown to generate identical estimates.

The primary focus of this paper is to characterize the efficiency gains resulting from supplementing CDA with direct metering data. Both a synthetic example and a real situation based on an actual load study involving 400 households in the state of New South Wales, Australia are used to illustrate these gains.

2. A RANDOM COEFFICIENT CDA MODEL WITH DIRECT METERING

2.1 Basic model

The basic CDA model for electricity consumption (possibly annual, monthly, daily, or hourly) is of the form

$$(1) \qquad y_i = z_i'\phi + d_i'\gamma_i^+ \qquad i=1,\ldots,N$$

where y_i = electricity consumption of customer i,

z_i' = row vector of observations on l explanatory variables,

d_i' = row vector of observations on k appliance dummies, the first of which is always unity.

The typical assumption that the coefficients of the appliance dummies are fixed, is unrealistic. There are two important sources of variation:

(i) during any particular hour the intensity of use of a particular appliance will vary from household to household,

(ii) the dummies indicate only absence or presence of the appliance and do not allow for variations in size or capacity.

Following Fiebig et al. (1988), we assume that

(2) $\gamma_i^+ = \gamma + v_i$

where γ is a k×1 vector of non-stochastic mean response coefficients, and $v_i' = (v_{1i}, v_{2i}, \ldots, v_{ki})$ is a vector of random disturbances.

Notice that (1) is written without a separate disturbance term. This omission is deliberate, as a separate disturbance can not be distinguished from v_{1i}, the disturbance associated with the intercept.

At this stage it is appropriate to recognize that some care needs to be exercised in defining random coefficients for dummy variables. According to (2), all elements of γ_i^+ are random. However, on observing the realized sample values for the dummy variables, (here appliance holdings), it is possible to identify some elements as identically equal to zero. Now a modified version of (2) is appropriate, namely

(2') $\gamma_i^* = \Delta_i \gamma_i^+$

where Δ_i is a k dimensional diagonal matrix whose diagonal elements are the appliance dummies, which are zero or unity depending on the appliance holdings of the ith customer.

Because $d_i'\Delta_i = d_i'$, the combination of (1) with either (2) or (2') yields a model of the form

(3) $y_i = z_i'\phi + d_i'\gamma + u_i$

where

(4) $u_i = d_i'v_i$

Assuming,

(5) $E(v_i) = 0, \quad E(v_i v_i') = A, \quad E(v_i v_j') = 0 \text{ for } i \neq j,$

it follows that

(6) $E(u_i) = 0, \quad E(u_i^2) = d_i'Ad_i, \quad E(u_iu_j) = 0$ for $i \neq j$.

This is a variant of the Hildreth-Houck random coefficient model (RCM). In Fiebig, Bartels and Aigner (1988), A was assumed to be a diagonal matrix, which implies a heteroskedastic error variance of the form:

(7) $\sigma_i^2 \equiv E(u_i^2) = d_i'\alpha$

where $\alpha' = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ is a vector comprising the diagonal elements of A.

In obvious notation write (3) as

(8) $y_i = x_i'\beta + u_i$

and let the full error covariance matrix be Ω, which is a diagonal matrix with typical diagonal element given by (7). Now for known Ω the GLS estimator of β is given by

(9) $\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$

It is also possible to predict the individual random response vector. The predictor

(10) $\hat{\gamma}_i^* = \Delta_i\hat{\gamma} + Ad_i(d_i'\alpha)^{-1}(\ y_i - x_i'\hat{\beta}\ )$

is best linear unbiased; see Griffiths (1972). For our particular problem these best linear unbiased predictors (BLUPs) are of great interest. They represent predictions of actual customer end-use loads and as such can be used to develop distributions of end-use loads over individuals. Operational variants of the estimator in (9) and the predictor in (10) have been proposed by Fiebig, Bartels and Aigner (1988). See also Bartels and Fiebig (1990) for further discussion.

2.2 Incorporating direct metering

Conditional demand analysis arrives at estimates for the load contribution of different end-uses by statistically disaggregating the total household loads for a sample of households. This is an indirect approach and the estimates it generates are often imprecise. The RCM is likely to provide some efficiency gains relative to OLS procedures but it remains the case that there is considerable room for improvement in these estimates. One obvious alternative is to directly meter specific appliances for a subsample of households. Suppose meters were allocated to all households possessing the kth appliance. Treatment of the resultant direct metering data would be straightforward: the mean of the observed loads being the estimate of the average load for that particular end-use. For the remaining end-uses the CDA model would be estimated after having subtracted the appliance loads of the kth appliance from the totals for the relevant households, and, after omitting the associated appliance dummy variable.

Complete metering is typically not a cost-effective alternative but in some cases a limited direct metering program may be feasible. Consequently we need to consider an appropriate method of incorporating direct metering data into the CDA framework.

The suggested approach follows directly from our random coefficient framework. Suppose direct metering information is available on the kth appliance for a total of n households where n is less than the number of households in our sample who have this appliance. For these households we observe a realization of the random response coefficient. Again this load can be subtracted from the household's total observed load and for these observations the appliance dummy set to zero. Finally, these adjusted observations are augmented to include the additional n observations that constitute the actual response coefficients of the kth appliance dummy. The stacked regression allows joint estimation of the mean response associated with the kth appliance, utilizing the data from households that were and were not directly metered.

The assumption that the covariance matrix for v_i is diagonal ensures that the error covariance matrix for the stacked regression is also diagonal with a heteroskedastic structure of the form discussed previously. In fact, it is as if there is an additional sample of n households with only one appliance. Notice also, that in the limiting case where every household with the kth appliance is directly metered, there is no gain from joint estimation and therefore this procedure reduces to that suggested for complete metering data.

2.3 A general model and alternative interpretation

Suppose direct metering information is available for p appliances and that these data were recorded for only a subset of the customers who have the respective appliances. Also, any single customer has at most one metered appliance. Consider two models, Model I being the basic CDA model and Model II its generalization that incorporates the direct metering data. For an arbitrary hour, these can be written in matrix form as:

$$(11) \qquad y_h = X_h \beta + u_h \qquad h = I, II$$

where

$$(12) \qquad y_I = \begin{bmatrix} y_a \\ y_b + y_c \end{bmatrix}, \quad y_{II} = \begin{bmatrix} y_a \\ y_b \\ y_c \end{bmatrix}$$

and

$$(13) \qquad u_I = \begin{bmatrix} u_a \\ u_b + u_c \end{bmatrix}, \quad u_{II} = \begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix}$$

while the design matrices for these models can be written as

$$(14) \qquad X_I = \left[\begin{array}{cccc} Z_0 & d_{01} & \cdots & d_{0p} \\ \hline Z_1 & d_{11} & \cdots & d_{1p} \\ \vdots & \vdots & & \vdots \\ Z_p & d_{p1} & \cdots & d_{pp} \end{array} \right]$$

and

$$(15) \qquad X_{II} = \left[\begin{array}{ccccc} Z_0 & d_{01} & d_{02} & \cdots & d_{0p} \\ \hline Z_1 & 0 & d_{12} & \cdots & d_{1p} \\ Z_2 & d_{21} & 0 & \cdots & d_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ Z_p & d_{p1} & d_{p2} & \cdots & 0 \\ \hline 0 & \iota_1 & 0 & \cdots & 0 \\ 0 & 0 & \iota_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \iota_p \end{array}\right]$$

where ι_j is an n_j column vector of unit elements. The top block of observations refers to the N-n ($n = \Sigma n_j$) customers who were not metered while the n customers who were metered have been arranged into groups according to the metered appliance to form the second block. For Model II there is a third block of n observations representing the observations on the directly metered appliances. In obvious notation these two design matrices can be written as

$$(16) \qquad X_I = \begin{bmatrix} X_a \\ X_b + X_c \end{bmatrix}, \quad X_{II} = \begin{bmatrix} X_a \\ X_b \\ X_c \end{bmatrix}$$

In a similar fashion the variance covariance matrices of the disturbance vectors are given by

$$(17) \qquad Var(u_I) = \Sigma_I = \begin{bmatrix} \Sigma_a & 0 \\ 0 & \Sigma_b + \Sigma_c \end{bmatrix}$$

$$(18) \qquad Var\ (u_{II}) = \Sigma_{II} = \begin{bmatrix} \Sigma_a & 0 & 0 \\ 0 & \Sigma_b & 0 \\ 0 & 0 & \Sigma_c \end{bmatrix}$$

where Σ_h, h = a,b,c are diagonal matrices.

While we have set up the model so that meters are assumed to be attached to individual appliances our framework could just as easily apply to the situation where meters recorded the load of a group of appliances. It could be that certain appliances are often put on the same circuit. In this case there would be p groups of appliances and the d_{ij} and ι_j vectors would simply become matrices with the number of columns equal to the number of appliances in the jth group.

Aigner and Schönfeld (1988) and Caves, Herriges, Train and Windle (1987) have recently suggested that limited direct metering information could be incorporated into CDA by treating it as stochastic prior information. Formally, define such a model, say Model III, by

(19) $$y_{III} = X_{III}\beta + u_{III}$$

where

$$y_{III} = \begin{bmatrix} y_a \\ y_b + y_c \\ y_c \end{bmatrix}, \quad X_{III} = \begin{bmatrix} X_a \\ X_b + X_c \\ X_c \end{bmatrix} \text{ and } u_{III} = \begin{bmatrix} u_a \\ u_b + u_c \\ u_c \end{bmatrix}$$

Notice that as in Model II, the basic data are augmented by the direct metering information, but unlike Model II, there is no adjustment of total loads or appliance dummies. The other difference is that the disturbance covariance matrix is no longer diagonal as the observations in the second and third blocks are correlated. The question of the comparison of the two approaches embodied in Models II and III is answered by the following proposition:

Proposition A: The GLS estimator associated with model II is identical to that associated with model III.

The proof is supplied in the Appendix.

Notice that this equivalence refers to a particular method of handling stochastic prior information. The result is unlikely to hold for alternatives that could come from a richer Bayesian framework.

3. EFFICIENCY GAINS FROM DIRECT METERING

3.1 Direct metering implies efficiency gains

An immediate question of interest is the characterization of the efficiency gains that arise from incorporating direct metering data, relative to the alternative of simply ignoring the information. In particular, this involves a comparison of the relative efficiency of the GLS estimators of Models I and II.

Proposition B: The GLS estimator associated with model II is more efficient than the estimator associated with model I. Denoting the variance-covariance matrices of the two estimators by V_I and V_{II} it is the case that:

(i) the difference between the variance-covariance matrices, $V_I - V_{II}$, is positive semi-definite (psd),

(ii) the trace of this difference, $tr(V_I - V_{II})$, is strictly positive.

The proof is supplied in the Appendix.

Different configurations of meters imply different forms for the X matrix and the error covariance matrix which translates into different covariance matrices for the β estimates. Given an actual data set and values for the variances of the random responses these covariance matrices can be determined and compared. We now do this for two particular models in order to further characterize the gains from direct metering.

3.2 Efficiency gains in a simple model

In order to further characterize the gain from direct metering, consider a simple CDA model comprising an intercept and two appliance dummies where direct metering is available for the first appliance. The model incorporating the direct metering information is given by,

$$(20) \qquad y_{II} = \begin{bmatrix} \iota_1 & \iota_1 & \iota_1 \\ \iota_2 & \iota_2 & 0 \\ \iota_3 & 0 & \iota_3 \\ \iota_4 & 0 & 0 \\ 0 & \iota_5 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + u_{II}$$

where there are a total of N_1 households that own both appliances, N_2

households that own only the first appliance, N_3 households that own only the second appliance, and N_4 households that do not own either appliance. There are a total of n households that have the first appliance metered, n_1 having been taken from the first group of households and the remaining n_2 from the second group. Consequently, $\iota_1, \ldots, \iota_5$ are unit vectors with dimensions: $N_1 - n_1$, $N_2 - n_2$, $N_3 + n_1$, $N_4 + n_2$ and n. The disturbance variance-covariance matrix is given by,

$$\text{(21)} \qquad \mathrm{Var}(u_{II}) = \mathrm{diag}(\sigma_1^2 I_1, \sigma_2^2 I_2, \sigma_3^2 I_3, \sigma_4^2 I_4, \sigma_5^2 I_5)$$

where $\sigma_1^2 = \alpha_1 + \alpha_2 + \alpha_3$, $\sigma_2^2 = \alpha_1 + \alpha_2$, $\sigma_3^2 = \alpha_1 + \alpha_3$, $\sigma_4^2 = \alpha_1$, and $\sigma_5^2 = \alpha_2$.

The first estimator of this model, denoted by $\hat{\beta}_I$, does not incorporate the direct metering information while the second, denoted by $\hat{\beta}_{II}$, does. Two experiments were conducted in order to compare the efficiencies of these two estimators the designs of which are given as:

Experiment 1: $N_1 = 200$, $N_2 = 100$, $N_3 = 50$, $N_4 = 50$, $n = 50$
$\alpha_1 = \alpha_3 = 1$, $\alpha_2 = 0.04, 0.25, 1, 4, 25$
$n_1 = 0, 10, 25, 40, 50$

Experiment 2: $N_1 = 200$, $N_2 = 100$, $N_3 = 50$, $N_4 = 50$
$\alpha_1 = \alpha_3 = 1$, $\alpha_2 = 0.04, 0.25, 1, 4, 25$
$n_1 = n = 10, 20, \ldots, 100$

Both experiments allow for a range of values for the relative variability of the random response of the metered appliance. In Experiment 1, this is coupled with variation in the composition of the type of household (according to appliance holdings) that is metered, while Experiment 2 varies the number of metered households.

For the purposes of comparison, the criterion chosen was the relative traces of the covariance matrices of $\hat{\beta}_{II}$ and $\hat{\beta}_I$. These quantities will be less than unity, with smaller values indicating greater efficiency gains from direct metering.

Results for Experiment 1 are given in Table 1. Numerically, these efficiency gains can be quite dramatic and display a strong inverse relationship with

α_2. While the effect of varying n_1 is small, notice the systematic tendency for the preferred value of n_1 to increase with larger α_2 values. In other words, the efficiency gains from direct metering are greatest when the variance of the random response of the metered appliance is relatively small, in which case there are small gains to be made from metering households with fewer appliances and hence smaller disturbance variances.

Results for Experiment 2 are given in Table 2. Again these efficiency gains can be quite dramatic. While efficiency gains are directly related to n, the number of meters, it is very interesting to note the considerable decrease in the rate of gain from additional meters especially when α_2 is small. In other words we rapidly reach a point where the gains from additional meters are likely to be outweighed by their costs. As an alternative way of viewing this phenomenon, consider $\alpha_2 = 1.0$. Here substantially more than 100 meters are required to obtain the type of efficiency gains provided by 10 meters when $\alpha_2 = 0.04$.

Table 1: Relative Efficiencies: Experiment 1

	α_2				
n_1	0.04	0.25	1.0	4	25
0	0.459	0.545	0.686	0.808	0.890
10	0.459	0.542	0.668	0.766	0.849
25	0.460	0.543	0.661	0.738	0.820
40	0.462	0.552	0.674	0.746	0.821
50	0.464	0.562	0.696	0.773	0.840

Table 2: Relative Efficiencies: Experiment 2

	α_2				
n	0.04	0.25	1.0	4	25
10	0.538	0.751	0.882	0.926	0.948
20	0.494	0.659	0.809	0.872	0.910
30	0.478	0.611	0.760	0.831	0.881
40	0.469	0.582	0.723	0.799	0.858
50	0.464	0.562	0.696	0.773	0.840
100	0.453	0.516	0.621	0.696	0.785

3.3 Efficiency gains in a second model

For the second model we draw heavily on the application of Fiebig, Bartels and Aigner (1988). The data for this study were compiled as part of the Domestic End-Use study conducted for the state of New South Wales (N.S.W.) in Australia under the auspices of an Industry Working Group comprising representatives from the Electricity Commission of N.S.W., the N.S.W. Department of Energy, the Local Government Energy Association and the County Councils of Sydney, Prospect, Southern Riverina, Illawarra and Shortland.

Load data consisted of hourly integrated demands for each customer averaged over working days for the month of July 1986. The resultant 24 observations for each customer represent the household's average working day load profile for that month. A selection of nine appliance dummies was chosen. These, together with their estimated population penetration rates, are provided in the following list:

- FREEZ = separate freezer (47%),
- FRIGAUT = automatic defrost fridge (53%),
- COOK = electric oven or hotplates (73%),
- DSH = dishwasher (22%),
- DRYER = clothes dryer (52%),
- HEAT = electric main or secondary heating (79%),
- HWPK = main tariff water heater (32%),
- HWOP = offpeak tariff water heater (51%),
- POOLPUMP = pool pump (6%).

The total sample size was 348. Direct metering information was available for two appliances, namely, HWOP and HWPK.

Empirical results from this study indicate substantial efficiency gains from the use of a random coefficient model and from the inclusion of directly metered observations. The improvements attributable to the latter were especially noteworthy prompting Fiebig et al. (1988, p.22) to recommend that: "... in future residential load studies every effort should be made to record loads of all appliances such as offpeak water heaters or ranges which are on a separate circuit running from the main board."

While their study highlights the potential gains from direct metering it provides little guidance on the appropriate allocation of meters across households and appliances. The actual allocation in the study was ad hoc. Would it have been better to allocate the available meters to ranges or some other mix of appliances? Given the actual data from this study and assuming the unknown variances to be given by their estimated values we are able to investigate such questions. These variances are not re-estimated for different configurations of meters.

In all our experiments we only consider those appliances most likely to be metered; namely COOK, HWPK and HWOP. Again the basic criteria used in evaluating alternative configurations of meters are the relative traces of the coefficient variance-covariance matrices. Because the estimation of the coefficients of the appliance dummies is our main concern these are only partial traces taken over the appliance coefficients and including the intercept.

Initially assume that each appliance is metered separately. Further suppose each appliance is completely metered; in other words all households possessing the appliance are assumed to be metered. The efficiency gains relative to the base of no metering are presented in Figure 1. Complete metering of any one of the three appliances would provide substantial efficiency gains throughout the day. No one appliance dominates the others throughout the day although metering HWOP provides the most gain for hours 1-15, 20-21 and 23-24, while in the remaining hours COOK is the best. However, COOK has 265 meters, HWOP 189 meters and HWPK only 105 meters. In fact the gains from metering HWPK seem qualitatively similar for say the hours 7-22 and yet involve substantially fewer meters and hence lower metering costs.

In order to focus our attention let us concentrate on hour 19. The results from Figure 1 for this hour are reproduced in Table 3, together with a series of other configurations of meters.

Figure 1:

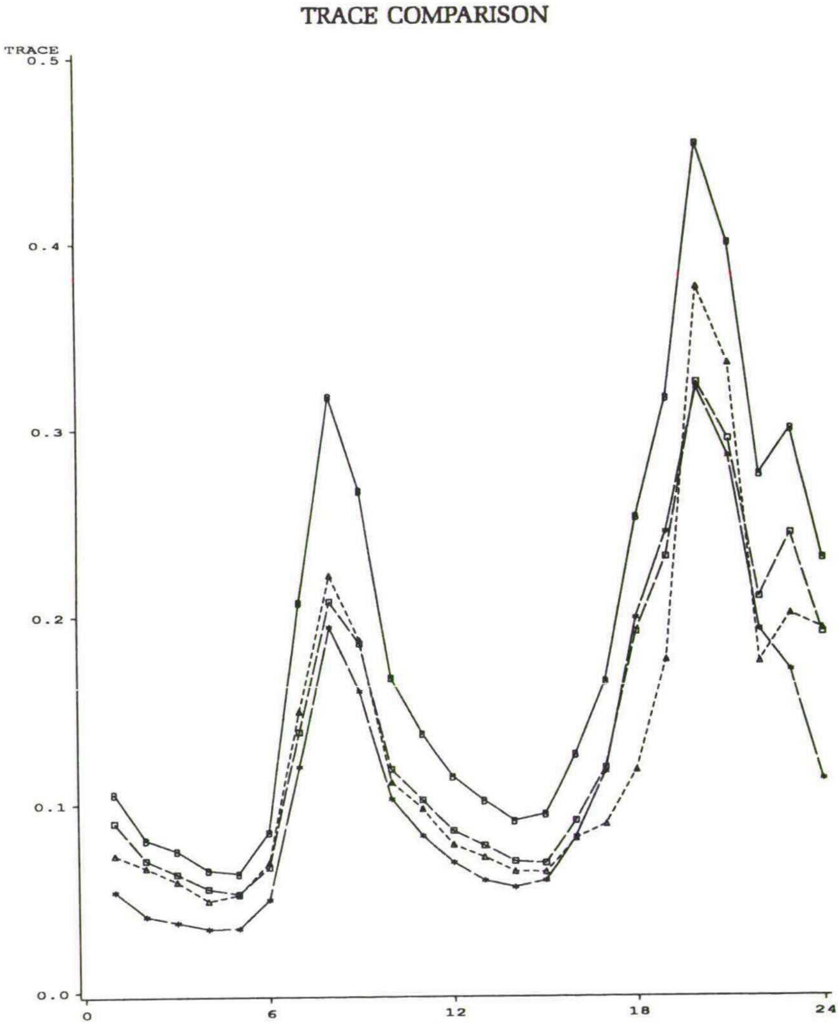
TRACE COMPARISON
TRACE
0.5
0.4
0.3
0.2
0.1
0.0
0
6
12
18
24
HOUR
B=BASE TRIANGLE=COOK SQUARE=HWPK STAR=HWOP

Table 3: Efficiency gains for hour 19

Meters			Partial trace	Ratio to base of no metering
COOK	HWOP	HWPK		
0	0	0	0.318	1.000
265	0	0	0.177	0.558
0	189	0	0.245	0.773
0	0	105	0.232	0.729
90	0	0	0.251	0.792
0	0	90	0.242	0.763
45	0	45	0.232	0.729
0	125	21	0.217	0.686

Notice that metering HWPK is unambiguously preferred to metering HWOP; it produces more efficient estimates with substantially fewer meters. This is a somewhat surprising result in view of the fact that the random coefficient variances are 0.118 for HWOP and 0.573 for HWPK. Results from the simple model of section 3.2 suggest that metering of appliances with lower variances is preferable. Naturally, this comparison does not control for other factors such as differences in the type of households. Apparently these factors have moderated the influence of the differences in variances.

For the comparison between COOK and HWPK the situation is unclear and an attempt was made to control for the differences in the number of meters. There is a group of 90 households that possess both COOK and HWPK. By metering households within this group a comparison can be made that controls for the number of meters and household characteristics. Again metering of HWPK is preferred. Because the random coefficient variance of COOK is 0.828 this result is consistent with the strategy of metering the lower variance appliance.

The efficiency comparisons have been limited to metering single appliances. Given the diminishing returns that were evident in the simple model of section 3.2 it is important to consider the possibility of metering more than one appliance.

Fiebig et al. (1988) report results where meters were available for HWOP and HWPK. A total of 125 out of the 189 households owning HWOP were directly metered while it was 21 out of 105 for HWPK. Results from this configuration of meters and a second example using the 90 households possessing both COOK and HWPK are presented in Table 3. For this second example the households were arbitrarily divided into two groups: the first 45 households were assumed to have COOK metered while the remaining 45 have HWPK metered.

It is clear that spreading the meters over different appliances is a good strategy. The 146 meters spread over the two hot-water appliances produces more efficient results than the 189 employed to totally meter HWOP. Similarly, spreading 90 meters between COOK and HWPK is preferred to locating them solely to only one of these appliances.

4. CONCLUSION

Imprecise end-use load estimates have been a major problem associated with conditional demand analysis. The use of direct metering information is one possible method of improving this situation. As the cost of direct metering comes down this approach will become increasingly attractive. The random coefficient formulation discussed here, provides a simple and intuitively appealing framework for the incorporation of limited direct metering data into conditional demand analysis. Importantly, it has been illustrated using both an artificial example and a real-life application that the efficiency gains from limited direct metering can be quite substantial.

Our analysis has provided some interesting insights into the problem of where, and, how many meters, need to be employed as part of a conditional demand study. In particular it seems that quite substantial gains may be achieved with only a relatively small number of meters. Given a choice between appliances it seems preferable to meter appliances for which the variation in use is smallest. Because of substantial diminishing returns when metering a single appliance, it is also advisable to spread the meters over different types of appliances.

Some of these optimal experimental design questions have been addressed in Aigner and Schönfeld (1988). Their analysis is limited, however, to the situation where households can only have two types of appliances. It is unlikely that useful algebraic results can be derived for the more general case of, say, 8 or 10 appliance types. A computational approach to the problem of how best to allocate a given number of direct metering devices is easily specified but takes on horrendous combinatorial dimensions. In practice, it is likely that the most fruitful approach will be to follow the example in this paper, and compare the efficiency gains from different feasible allocation schemes. Done systematically this could be likened to an heuristic optimization scheme. It is fair to conclude that a much more comprehensive study of the optimal placement of meters remains to be undertaken.

Appendix: Proofs of Propositions

Proof of Proposition A:

The proof follows directly from noting that Model III can be rewritten as

(A.1) $$B\, y_{II} = B\, X_{II}\beta + B\, u_{II}$$

where $B = \begin{bmatrix} I_{N-n} & 0 & 0 \\ 0 & I_n & I_n \\ 0 & 0 & I_n \end{bmatrix}$ and $\mathrm{Var}\,(u_{III}) = B\,\Sigma_{II} B\,'$.

Hence,

$$\hat{\beta}_{III} = [X_{II}'B\,'(B\,\Sigma_{II}B\,')^{-1}B\,X_{II}]^{-1}X_{II}'B\,'(B\,\Sigma_{II}B\,')^{-1}B\,y_{II}$$
$$= [X_{II}'\Sigma_{II}^{-1}X_{II}]^{-1}X_{II}'\Sigma_{II}^{-1}\,y_{II} = \hat{\beta}_{II} \quad \blacksquare$$

Proof of Proposition B:

The following lemmas are useful in the proof:

Lemma 1: Let S be a positive definite, symmetric matrix of rank (N+n), and R be an arbitrary matrix of dimension $N\times(N+n)$ and rank N, then $C = S^{-1} - R'(RSR')^{-1}R$ is positive semi-definite.

Proof of Lemma 1:

$C = S^{-1} - R'(RSR')^{-1}R$ is psd

$\Leftrightarrow \quad S^{-1/2}[I - S^{1/2}R'(RSR')^{-1}RS^{1/2}]S^{-1/2}$ is psd

$\Leftrightarrow \quad B = I - S^{1/2}R'(RSR')^{-1}RS^{1/2}$ is psd

But B is idempotent and hence psd $\blacksquare$

Lemma 2: Let X be an arbitrary matrix of dimension $(N+n)\times k$ and rank k, then the ranks of X'CX and CX are equal and X'CX = 0 iff CX = 0.

Proof of Lemma 2:

Define $X^* = S^{-1/2}X$ then $X'CX = X^{*\prime}BX^* = (BX^*)'(BX^*)$ since B is symmetric idempotent. Now $r[(BX^*)'(BX^*)] = r(BX^*)$ and $(BX^*)'(BX^*) = 0$ iff $BX^* = 0$.

But $CX = S^{-1/2}BX^*$ and the same results apply to CX as $S^{-1/2}$ is nonsingular.

$\blacksquare$

Now note that $X_I = AX_{II}$ and $\Sigma_I = A\Sigma_{II}A'$ where $A = \begin{bmatrix} I_{N-n} & 0 & 0 \\ 0 & I_n & I_n \end{bmatrix}$.

Hence the GLS estimators associated with Models I and II have precision matrices given by

(A.2) $$P_I = V_I^{-1} = X_I'\Sigma_I^{-1}X_I = X_{II}'A'(A\Sigma_{II}A')^{-1}AX_{II}$$

and

(A.3) $$P_{II} = V_{II}^{-1} = X_{II}'\Sigma_{II}^{-1}X_{II}$$

From (A.2) and (A.3) we have

(A.4) $$P_{II} - P_I = X_{II}'(\Sigma_{II}^{-1} - A'(A\Sigma_{II}A')^{-1}A)X_{II} = X_{II}'CX_{II}$$

By Lemma 1, C is psd and hence $X_{II}'CX_{II}$ is psd. Now $V_I - V_{II} = P_I^{-1} - P_{II}^{-1}$ is pd, psd or zero iff $P_{II} - P_I$ is pd, psd or zero.

Part (ii) of the Proposition follows if $X_{II}'CX_{II}$ is also nonzero. Note that C has the following structure

$$C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Sigma_b^{-1}-(\Sigma_b+\Sigma_c)^{-1} & -(\Sigma_b+\Sigma_c)^{-1} \\ 0 & -(\Sigma_b+\Sigma_c)^{-1} & \Sigma_c^{-1}-(\Sigma_b+\Sigma_c)^{-1} \end{bmatrix}$$

and thus

$$CX_{II} = \begin{bmatrix} 0 \\ [\Sigma_b^{-1}-(\Sigma_b+\Sigma_c)^{-1}]X_b - (\Sigma_b+\Sigma_c)^{-1}X_c \\ -(\Sigma_b+\Sigma_c)^{-1}X_b + [\Sigma_c^{-1}-(\Sigma_b+\Sigma_c)^{-1}]X_c \end{bmatrix}$$

Since all the Σ's are diagonal and X_c has unit elements where X_b has zeroes, it follows that $CX_{II} \neq 0$ implying that $X_{II}'CX_{II} \neq 0$ and hence Lemma 2 applies. ∎

REFERENCES

Aigner, D.J., and Schönfeld, P. (1988). "Experimental design for direct metering of residential electricity end-uses", University of Bonn, Econometrics Discussion Papers, #B-109.

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

Bartels, R. and Fiebig, D.G. (1990). "Constrained estimation of the Hildreth-Houck random coefficient model", University of Sydney, Econometrics Discussion Papers, #90-02.

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.

Fiebig, D.G., Bartels, R. and Aigner, D.J. (1988). "A random coefficient approach to the estimation of residential end-use load profiles", forthcoming Journal of Econometrics.

Griffiths, W.E. (1972). "Estimation of actual response coefficients in the Hildreth-Houck random coefficients model", Journal of the American Statistical Association 67, 633-635.

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.

Huss, W. (1985). "What makes a good load forecast?", Public Utilities Fortnightly, 116 (11), 27-35.

Parti, M. and Parti, C. (1980). "The total and appliance specific conditional demand for electricity in the household sector", Bell Journal of Economics, 11, 309-324.

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
8935	C. Dang and D. Talman	A New Triangulation of the Unit Simplex for Computing Economic Equilibria
8936	Th. Nijman and M. Verbeek	The Nonresponse Bias in the Analysis of the Determinants of Total Annual Expenditures of Households Based on Panel Data
8937	A.P. Barten	The Estimation of Mixed Demand Systems
8938	G. Marini	Monetary Shocks and the Nominal Interest Rate
8939	W. Güth and E. van Damme	Equilibrium Selection in the Spence Signaling Game
8940	G. Marini and P. Scaramozzino	Monopolistic Competition, Expected Inflation and Contract Length
8941	J.K. Dagsvik	The Generalized Extreme Value Random Utility Model for Continuous Choice
8942	M.F.J. Steel	Weak Exogenity in Misspecified Sequential Models
8943	A. Roell	Dual Capacity Trading and the Quality of the Market
8944	C. Hsiao	Identification and Estimation of Dichotomous Latent Variables Models Using Panel Data
8945	R.P. Gilles	Equilibrium in a Pure Exchange Economy with an Arbitrary Communication Structure
8946	W.B. MacLeod and J.M. Malcomson	Efficient Specific Investments, Incomplete Contracts, and the Role of Market Alternatives
8947	A. van Soest and A. Kapteyn	The Impact of Minimum Wage Regulations on Employment and the Wage Rate Distribution
8948	P. Kooreman and B. Melenberg	Maximum Score Estimation in the Ordered Response Model
8949	C. Dang	The D_3-Triangulation for Simplicial Deformation Algorithms for Computing Solutions of Nonlinear Equations
8950	M. Cripps	Dealer Behaviour and Price Volatility in Asset Markets
8951	T. Wansbeek and A. Kapteyn	Simple Estimators for Dynamic Panel Data Models with Errors in Variables

No.	Author(s)	Title
8952	Y. Dai, G. van der Laan, D. Talman and Y. Yamamoto	A Simplicial Algorithm for the Nonlinear Stationary Point Problem on an Unbounded Polyhedron
8953	F. van der Ploeg	Risk Aversion, Intertemporal Substitution and Consumption: The CARA-LQ Problem
8954	A. Kapteyn, S. van de Geer, H. van de Stadt and T. Wansbeek	Interdependent Preferences: An Econometric Analysis
8955	L. Zou	Ownership Structure and Efficiency: An Incentive Mechanism Approach
8956	P.Kooreman and A. Kapteyn	On the Empirical Implementation of Some Game Theoretic Models of Household Labor Supply
8957	E. van Damme	Signaling and Forward Induction in a Market Entry Context
9001	A. van Soest, P. Kooreman and A. Kapteyn	Coherency and Regularity of Demand Systems with Equality and Inequality Constraints
9002	J.R. Magnus and B. Pesaran	Forecasting, Misspecification and Unit Roots: The Case of AR(1) Versus ARMA(1,1)
9003	J. Driffill and C. Schultz	Wage Setting and Stabilization Policy in a Game with Renegotiation
9004	M. McAleer, M.H. Pesaran and A. Bera	Alternative Approaches to Testing Non-Nested Models with Autocorrelated Disturbances: An Application to Models of U.S. Unemployment
9005	Th. ten Raa and M.F.J. Steel	A Stochastic Analysis of an Input-Output Model: Comment
9006	M. McAleer and C.R. McKenzie	Keynesian and New Classical Models of Unemployment Revisited
9007	J. Osiewalski and M.F.J. Steel	Semi-Conjugate Prior Densities in Multi-variate t Regression Models
9008	G.W. Imbens	Duration Models with Time-Varying Coefficients
9009	G.W. Imbens	An Efficient Method of Moments Estimator for Discrete Choice Models with Choice-Based Sampling
9010	P. Deschamps	Expectations and Intertemporal Separability in an Empirical Model of Consumption and Investment under Uncertainty

No.	Author(s)	Title
9011	W. Güth and E. van Damme	Gorby Games - A Game Theoretic Analysis of Disarmament Campaigns and the Defense Efficiency-Hypothesis
9012	A. Horsley and A. Wrobel	The Existence of an Equilibrium Density for Marginal Cost Prices, and the Solution to the Shifting-Peak Problem
9013	A. Horsley and A. Wrobel	The Closedness of the Free-Disposal Hull of a Production Set
9014	A. Horsley and A. Wrobel	The Continuity of the Equilibrium Price Density: The Case of Symmetric Joint Costs, and a Solution to the Shifting-Pattern Problem
9015	A. van den Elzen, G. van der Laan and D. Talman	An Adjustment Process for an Exchange Economy with Linear Production Technologies
9016	P. Deschamps	On Fractional Demand Systems and Budget Share Positivity
9017	B.J. Christensen and N.M. Kiefer	The Exact Likelihood Function for an Empirical Job Search Model
9018	M. Verbeek and Th. Nijman	Testing for Selectivity Bias in Panel Data Models
9019	J.R. Magnus and B. Pesaran	Evaluation of Moments of Ratios of Quadratic Forms in Normal Variables and Related Statistics
9020	A. Robson	Status, the Distribution of Wealth, Social and Private Attitudes to Risk
9021	J.R. Magnus and B. Pesaran	Evaluation of Moments of Quadratic Forms in Normal Variables
9022	K. Kamiya and D. Talman	Linear Stationary Point Problems
9023	W. Emons	Good Times, Bad Times, and Vertical Upstream Integration
9024	C. Dang	The D_2-Triangulation for Simplicial Homotopy Algorithms for Computing Solutions of Nonlinear Equations
9025	K. Kamiya and D. Talman	Variable Dimension Simplicial Algorithm for Balanced Games

No.	Author(s)	Title
9026	P. Skott	Efficiency Wages, Mark-Up Pricing and Effective Demand
9027	C. Dang and D. Talman	The D_1-Triangulation in Simplicial Variable Dimension Algorithms for Computing Solutions of Nonlinear Equations
9028	J. Bai, A.J. Jakeman and M. McAleer	Discrimination Between Nested Two- and Three-Parameter Distributions: An Application to Models of Air Pollution
9029	Th. van de Klundert	Crowding out and the Wealth of Nations
9030	Th. van de Klundert and R. Gradus	Optimal Government Debt under Distortionary Taxation
9031	A. Weber	The Credibility of Monetary Target Announcements: An Empirical Evaluation
9032	J. Osiewalski and M. Steel	Robust Bayesian Inference in Elliptical Regression Models
9033	C. R. Wichers	The Linear-Algebraic Structure of Least Squares
9034	C. de Vries	On the Relation between GARCH and Stable Processes
9035	M.R. Baye, D.W. Jansen and Q. Li	Aggregation and the "Random Objective" Justification for Disturbances in Complete Demand Systems
9036	J. Driffill	The Term Structure of Interest Rates: Structural Stability and Macroeconomic Policy Changes in the UK
9037	F. van der Ploeg	Budgetary Aspects of Economic and Monetary Integration in Europe
9038	A. Robson	Existence of Nash Equilibrium in Mixed Strategies for Games where Payoffs Need not Be Continuous in Pure Strategies
9039	A. Robson	An "Informationally Robust Equilibrium" for Two-Person Nonzero-Sum Games
9040	M.R. Baye, G. Tian and J. Zhou	The Existence of Pure-Strategy Nash Equilibrium in Games with Payoffs that are not Quasiconcave
9041	M. Burnovsky and I. Zang	"Costless" Indirect Regulation of Monopolies with Substantial Entry Cost
9042	P.J. Deschamps	Joint Tests for Regularity and Autocorrelation in Allocation Systems

No.	Author(s)	Title
9043	S. Chib, J. Osiewalski and M. Steel	Posterior Inference on the Degrees of Freedom Parameter in Multivariate-t Regression Models
9044	H.A. Keuzenkamp	The Probability Approach in Economic Methodology: On the Relation between Haavelmo's Legacy and the Methodology of Economics
9045	I.M. Bomze and E.E.C. van Damme	A Dynamical Characterization of Evolutionarily Stable States
9046	E. van Damme	On Dominance Solvable Games and Equilibrium Selection Theories
9047	J. Driffill	Changes in Regime and the Term Structure: A Note
9048	A.J.J. Talman	General Equilibrium Programming
9049	H.A. Keuzenkamp and F. van der Ploeg	Saving, Investment, Government Finance and the Current Account: The Dutch Experience
9050	C. Dang and A.J.J. Talman	The D_1-Triangulation in Simplicial Variable Dimension Algorithms on the Unit Simplex for Computing Fixed Points
9051	M. Baye, D. Kovenock and C. de Vries	The All-Pay Auction with Complete Information
9052	H. Carlsson and E. van Damme	Global Games and Equilibrium Selection
9053	M. Baye and D. Kovenock	How to Sell a Pickup Truck: "Beat-or-Pay" Advertisements as Facilitating Devices
9054	Th. van de Klundert	The Ultimate Consequences of the New Growth Theory; An Introduction to the Views of M. Fitzgerald Scott
9055	P. Kooreman	Nonparametric Bounds on the Regression Coefficients when an Explanatory Variable is Categorized
9056	R. Bartels and D.G. Fiebig	Integrating Direct Metering and Conditional Demand Analysis for Estimating End-Use Loads

P.O. BOX 90153, 5000 LE TILBURG, THE NETHERLANDS