# RESIDENTIAL ELECTRICITY DEMAND UNDER TIME-OF-DAY PRICING: EXPLORATORY DATA ANALYSIS FROM THE LOS ANGELES RATE STUDY

PREPARED FOR THE LOS ANGELES DEPARTMENT OF WATER AND POWER
AND THE JOHN A. HARTFORD FOUNDATION

WILLARD G. MANNING, JR., AND JAN PAUL ACTON

R-2426-DWP/HF
DECEMBER 1980

This report has been prepared as part of a project for the Los Angeles Department of Water and Power to study electricity rate structures (under Contract No. 10189). The project is funded in part by a cooperative agreement between the City of Los Angeles and the Department of Energy. Additional support was provided by a grant from the John A. Hartford Foundation.

The Rand Publications Series: The Report is the principal publication documenting and transmitting Rand's major research findings and final research results. The Rand Note reports other outputs of sponsored research for general distribution. Publications of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

Published by The Rand Corporation

# RESIDENTIAL ELECTRICITY DEMAND UNDER TIME-OF-DAY PRICING: EXPLORATORY DATA ANALYSIS FROM THE LOS ANGELES RATE STUDY

PREPARED FOR THE LOS ANGELES DEPARTMENT OF WATER AND POWER
AND THE JOHN A. HARTFORD FOUNDATION

WILLARD G. MANNING, JR., AND JAN PAUL ACTON

R-2426-DWP/HF
DECEMBER 1980

Rand
SANTA MONICA, CA. 90406

# PREFACE

The Los Angeles Electricity Rate Study--a five-year project concluded in 1980--was designed to yield information about the effects of alternative pricing structures for residential consumers. The experimental plans included time-of-day, seasonal, and time-invariant rates for approximately 1800 households over a 30-month period. The study was conducted jointly by the Los Angeles Department of Water and Power and The Rand Corporation, with partial funding from the U.S. Department of Energy.

This report presents results of behavior under the time-of-use plans, especially those which charge a different price for electricity at certain hours. Such time-of-use rates are being actively considered in most regulatory jurisdictions, but there is considerable uncertainty about their effects among residential users. Since the Los Angeles experiment was designed to yield demand curve estimates, results of this study should be useful in forecasting likely effects in a number of different service territories.

This report concentrates on exploratory data analysis in order to understand the fundamental properties of the data and to obtain valid statistical tests. The resulting estimates are highly robust and the magnitudes appear reasonable. The models support analysis of electricity use in varied circumstances and permit forecasting consumption under new rate structures. The report should be of particular interest to electric utilities and regulatory bodies investigating consumer responses to electricity prices. For economists and policy analysts it provides a methodology for approaching the analysis of time-of-use pricing from a relatively "model-free," statistical perspective.

Related Rand research on electricity pricing and demand can be found in the following publications:

- *Design of the Los Angeles Peak-Load Pricing Experiment for Electricity*, Willard G. Manning, Jr., Bridger M. Mitchell, and Jan Paul Acton, R-1955-DWP, November 1976.

- *Lessons To Be Learned from the Los Angeles Rate Experiment in Electricity*, Jan Paul Acton, Willard G. Manning, Jr., and Bridger M. Mitchell, R-2113-DWP, July 1978.
- *Peak-Load Pricing: European Lessons for U.S. Energy Policy*, Bridger M. Mitchell, Willard G. Manning, Jr., and Jan Paul Acton, Ballinger Publishing Company, 1978.
- *The Los Angeles Senior Citizen Lifeline Electricity Rate*, Timothy J. Sullivan, R-2278-DWP/NSF, January 1979.
- *Sample Selection in the Los Angeles Electricity Rate Study*, Winston Chow and Bridger M. Mitchell, R-2430-DWP, July 1979.
- *Conducting a Survey Using the Client's Staff: Evaluation of Interviewer Performance in the Electricity Rate Study*, Sandra Berry, R-2223-DWP, September 1979.
- *Evaluating Time-of-Day Electricity Rates for Residential Customers*, Jan Paul Acton and Bridger M. Mitchell, R-2509-DWP, November 1979.
- *Do Time of Use Rates Change Load Curves?*, Jan Paul Acton and Bridger M. Mitchell, R-2588-DWP/EPRI, May 1980.
- *The Effect of Time-of-Day Rates in the Los Angeles Electricity Study*, Bridger M. Mitchell and Jan Paul Acton, N-1533-DWP, October 1980.
- *Planning, Processing, and Analyzing Data for Residential Load Studies*, Jan Paul Acton and Bridger M. Mitchell, N-1534-DWP, June 1980.
- *Seasonal Electricity Demand and Pricing: Analysis with a Variable Response Model*, Lee Lillard and Jan Paul Acton, R-2425-DWP, May 1980.
- *Electricity Consumption by Time of Use in a Hybrid Demand System*, Bridger M. Mitchell and Jan Paul Acton, R-2628-DWP/HF, December 1980.

## SUMMARY

The Los Angeles Electricity Rate Study is one of several electricity rate experiments conducted between 1975 and 1980. Because of its complex experimental design and extensive survey information, data from the Los Angeles study can be used to analyze a number of alternative time-of-day rates in a variety of situations. In many applications, it will be desirable to analyze the data using a system of economic demand curves that take advantage of the implications of utility maximizing behavior by customers in order to gain more powerful implications from the data.

Before examining the Los Angeles data with such highly structured models, we advocate exploratory data analysis with relatively "model free" statistical demand curves. In this exploratory phase, we place our emphasis on the statistical properties of the data in order to obtain efficient coefficients of the estimates and valid tests of statistical significance. This leads to relatively robust and accurate models which permit forecasts over a variety of circumstances. Once the fundamental properties of the data are understood, they can be used in more highly structured models where it may be very difficult to examine these more fundamental characteristics of the data.

We broke the analysis into two parts. First, we sought transformations of the dependent and explanatory variables using data that were averaged over an 18-month period. In this step, we examined in particular the distribution of the error terms in order to have valid statistical tests. Second, we estimated the full model with a pooled time series of cross-sectional data. We allowed for serial correlation by using a random effects variance components model, and tested for possible self-selection bias by households who participated in the rate experiment.

We found that a double logarithmic specification provides the best estimating equation by the criteria we selected. It is very unlikely that we would have selected such a specification a priori, but the

estimation results appeared to be very robust and the magnitudes are sensible. The own price elasticities range from about -.05 to -.10 on average across all households. These are similar magnitudes to ones estimated from the Wisconsin and Arizona rate experiments by a number of analysts. Cross prices tend to be statistically insignificant, although of the correct sign. Appliances, income, and other explanatory variables have the expected effects.

These average price elasticities mask a considerable variation among households, depending on levels of use and appliance holdings. When we test a more complex specification that has interactions between prices and appliances, we find important variation. Although individual coefficients on interaction terms are generally not statistically significantly different from zero, the test for dropping a vector of the interactions is statistically different from 0. When we select a model that is moderately complex, allowing for interactions between appliances and own prices (but not cross prices), we find that price elasticities increase systematically with level of use and appliances. Households with few appliances have generally insignificant elasticities between 0 and -.10. At the other extreme, households with extensive appliance holdings and high average monthly levels of use have elasticities between -.20 and -.40 during daytime hours. In particular, the presence of a swimming pool or air conditioner increases the own price elasticities significantly.

It lies beyond the scope of the present report to use these estimates for systematic forecasting and for benefit-cost analysis of alternative time-of-day rates. But the flexible, statistically significant coefficients that have resulted from this exploratory data analysis should be directly useful in a number of applications. Furthermore, once the exploratory data analysis is completed and the fundamental properties of the data are understood, we can proceed to estimate more highly structured models with some confidence that the properties of the transformations are known and the nature of the error structure is understood.

## ACKNOWLEDGMENTS

We are indebted to a number of people for extensive data preparation, computation, and statistical analysis that made this exploratory data analysis possible. We thank Carolyn Lee, Carol Medine, Mike Rabin, Mitch Tuller, and Ed Von Heydenreich for careful assistance and programming. Winston Chow, Carl Morris, and Dan Relles have provided valuable statistical advice. Dan Relles also provided the statistical software required for the time-series data and analysis. Winston Chow led us to examine the possibility of selection bias more closely, even when it appeared on the surface to be unimportant.

Emmett Keeler, Dan Kohler, Hethie Parmesano, and Dennis Whitney made useful comments on earlier drafts. Will Harriss's editorial skills improved the text considerably. Sandee Sims cheerfully took us through several redraftings of the paper.

We gratefully acknowledge the support of the Los Angeles Department of Water and Power and especially Michael Moore, the original project officer. A grant from the John A. Hartford Foundation is permitting us to broaden this analysis into a more national scope.

CONTENTS

PREFACE ..... iii

SUMMARY ..... v

ACKNOWLEDGMENTS ..... vii

Section
I. INTRODUCTION ..... 1
Our Approach and Main Findings ..... 2
Outline of the Report ..... 3

II. THE LOS ANGELES EXPERIMENT AND DATA BASE ..... 4
Overview of the TOD Experiment ..... 4
Other Data ..... 6
Unit of Analysis ..... 6

III. EXPLORATORY ANALYSIS ..... 8
Transformation of the Dependent Variable ..... 9
Transformations of the Explanatory Variables ..... 11
Analysis of Residuals and Outlier Observations ..... 17

IV. ESTIMATION RESULTS FOR POOLED TIME-SERIES AND CROSS-SECTION DATA ..... 19
Serial Correlation ..... 19
Testing for Self-Selection Bias ..... 23
Two Models of TOD Demand ..... 25
Directions for Future Analysis ..... 33

V. CONCLUSION ..... 35

REFERENCES ..... 37

# I. INTRODUCTION

The main purpose of this report is to urge analysts to apply exploratory data analysis with relatively simple, "model-free" specifications before analyzing time-of-day response with a more highly structured set of economic demand equations. We illustrate such an exploratory analysis with a large set of experimental data and report the estimates and statistical tests that result. We then compare our results with those obtained by some other researchers analyzing time-of-day electricity use.

The Los Angeles Electricity Rate Study is one of several time-of-day (TOD) experiments conducted between 1975 and 1980. Good empirical estimates of price elasticities from these experiments will be useful for a number of policy deliberations that require forecasts of the expected effects of TOD rates. For most rate deliberations (including those required by the PURPA[1] standards), policymakers must determine the effects of prices in different periods of time on the quantity and timing of electricity use, as well as the distribution of effects across customers. The answer to these policy questions is beyond the scope of the present report.[2] But the empirical results reported from an analysis of the Los Angeles experiment are a necessary ingredient of such deliberations.

The large amount of data collected during the five-year Los Angeles Electricity Rate Study permits a rich model specification--including models that take account of the implications of economic theory. Systems of demand equations, such as the translog or generalized Leontief system, are candidates for such applications, and they have been applied to some

---

[1] The Public Utilities Regulatory Policies Act of 1978 requires hearings by all state regulatory commissions and most larger publicly controlled utilities on 11 different ratemaking standards, including time-of-day and seasonal electricity rates.

[2] These require a comparison of benefits in terms of producer and consumer welfare, as well as the costs of implementing a more complex and costly pricing structure.

of the other experiment data sets.[3]

Before plunging into such highly structured models, we decided to perform systematic exploratory data analysis--a technique more familiar to statisticians than to economists. We believe it is useful with such a large data set as the present one for several reasons. First, until exploratory analysis has been performed, one cannot know how sensitive the results may be to the distribution of the data or the presence of outliers that may unduly influence the results. Second, to make valid inferences, one needs consistent estimates of test statistics. The standard statistical tests require normally distributed error terms. Third, highly structured models may be very sensitive to minor deviations from normality in the data set, but their complexity may make it very difficult to detect such effects in a complete system of equation estimates. Finally, even without highly structured models, we have an extremely large number of candidate explanatory variables, for which economic theory offers little guidance in the choice of variables and their appropriate specification. Exploratory data analysis can help determine the appropriate specification and, in many cases, indicate appropriate aggregation or deletion of some of these variables.

## OUR APPROACH AND MAIN FINDINGS

Rather than address all of these analytic issues simultaneously, we examined them sequentially. First, using a random sample of the data, we sought a suitable transformation of the dependent variable using Box-Cox techniques. We found that a logarithmic transformation produced a variable which is approximately normally distributed. Second, using data averaged over months, we sought appropriate specification of the explanatory variables: appliances, demographics, weather, and prices. Again, we found that a logarithmic transformation of most explanatory variables yielded a fairly homoskedastic error structure. Third, we used pooled monthly data on households to obtain efficient coefficient estimates. In this time-series analysis

---

[3] See, among others, Caves and Christensen (1980), analysis of the Wisconsin data, and Aigner and Hausman (1980) or Kohler (1980), analysis of the Arizona experimental data. See also Aigner and Poirier (1980) for a critique of systems of demand equations in electricity and a review of several studies.

we examine the decision to enroll in the study to see if there is evidence of a selection bias and we test for serial correlation in the monthly observations. We find that the estimates of price responsiveness are sensitive to the decision to participate and make a correction for self-selection bias. We also find evidence of strong positive serial correlation in the data and report estimates using a random effects variance components model. Finally, in estimating with pooled time-series data, we test for differential price response by appliance holding and find some evidence of difference, but the individual effects are not very precise.

## OUTLINE OF THE REPORT

Section II below briefly describes the experiment and data available from the Los Angeles Electricity Study. Section III describes the exploratory data analysis that takes place on average data in order to select the basic specification. Section IV reports the results of the full estimation using a random effects model and allowing for selection bias. In this section we pay particular attention to the price coefficients and the possibility of interactions between prices and certain explanatory variables.

## II. THE LOS ANGELES EXPERIMENT AND DATA BASE

The data used in this analysis are from the Los Angeles Electricity Rate Study, a voluntary social experiment in time-of-use pricing of electricity for residential customers. The study included approximately 2,000 households assigned to a variety of time-of-day, seasonal, and time-invariant electricity rates. The households were enrolled for a 30-month period, with the first families joining the study in 1976. In the present analysis, we examine only the effects of time-of-day (TOD) pricing; seasonal and time-invariant rates are examined in Lillard and Acton (1980).

### OVERVIEW OF THE TOD EXPERIMENT

Experimental prices were selected to provide price variation under a range of possible costs that utilities might face in the early 1980s, when the experimental results would be available.[1] Furthermore, the Los Angeles study was consciously designed to include rate structures and price levels that might be relevant to situations other than that of Los Angeles. The combinations of experimental rates permit identifying own- and cross-price effects in each of the five distinct pricing periods per day and detecting nonlinearities in each of the time periods. Peak-period prices ranged from 5¢ to 13¢ per kwh and off-peak prices from 1¢ to 2¢ per kwh. Seasonal and flat rates ranged from 2¢ to 8¢ per kwh. Peak price periods varied from as little as three hours per day to as long as twelve hours per day. A total of 17 TOD rates were used, in addition to seven time-invariant rates.

Table 1 shows the breakdown of experimental prices and the number of households with continuous recording cassette meters used for this analysis. The sample contains 1047 households on time-of-day tariffs and 101 households

---

[1] The experimental design extended procedures developed by Conlisk and Watts (1969). Individual households were assigned to experimental treatments to achieve optimal efficiency and representativeness according to procedures developed by Morris (1979). See Manning, Mitchell, and Acton (1979) for a thorough discussion of the design. The sampling characteristics are described in Chow and Mitchell (1979).

Table 1

NUMBER OF HOUSEHOLDS BY EXPERIMENTAL RATE USED IN THE ANALYSIS

| Peak Period | Off-peak Price (¢/kwh) | Peak Price (¢/kwh) | Number of Households |
|---|---|---|---|
| Time-invariant | N/A | 2 | 59 |
| | N/A | 5 | 42 |
| 9 a.m.-noon | 2 | 5 | 78 |
| | 2 | 9 | 60 |
| | 2 | 13 | 20 |
| Noon-3 p.m. | 2 | 9 | 38 |
| | 2 | 13 | 20 |
| 3 p.m.-6 p.m. | 2 | 5 | 20 |
| | 2 | 9 | 79 |
| | 2 | 13 | 75 |
| 6 p.m.-9 p.m. | 2 | 5 | 20 |
| | 2 | 9 | 56 |
| | 2 | 13 | 57 |
| 3 p.m.-9 p.m. | 2 | 7 | 41 |
| Noon-9 p.m. | 1 | 5 | 121 |
| | 1 | 9 | 60 |
| 9 a.m.-9 p.m. | 1 | 5 | 60 |
| | 1 | 9 | 81 |
| 9 p.m.-9 a.m. | 2 | 5 | 161 |
| | | Total | 1148 |

on time-invariant rates.[2] The results reported here are based on data from the pre-enrollment interview and weekday consumption registered on cassette recorders between July 1976 and December 1977.

## OTHER DATA

Demographic and household information were obtained through an in-person household survey conducted before enrollment in the experiment.[3] In that interview, we determined family composition, appliance holdings (for some 16 difference appliances, identified by fuel type), appliance use habits, income, and other demographic information.

Weather data were gathered hourly for three distinct weather zones in the service territory of the Los Angeles Department of Water and Power: a moderate coastal zone, an inland valley with more extreme temperatures in both summer and winter, and a civic center area with an intermediate climate. We represent the effects of temperature with heating degree hours and cooling degree hour variables each month.[4]

## UNIT OF ANALYSIS

The household is the unit of observation for this analysis. Data are available for a maximum of 18 months per household for this exploratory analysis. On average, however, households were observed for 13 months in this period of time because households were enrolled on a staggered basis beginning summer of 1976. The analysis is all conducted

---

[2] Table 1 gives the number of households used for this analysis. Because of early attrition and meter failures, the number differs for the enrollment targets given in Manning, Mitchell, and Acton (1979, p. 175).

[3] We also interviewed households at the mid-point and end of their participation in the experiment, but only the baseline survey information is used for this analysis.

[4] The number of cooling degrees in a one-hour period = max [temperature - 65°F, 0]. Heating degree hours are defined as max [65°F - temperature, 0]. The heating degree hour and cooling degree hour variables capture the fluctuations in temperatures faced by a household in much greater detail than do the more conventionally used heating degree day variables, which are based on average daily values of temperature.

in terms of average monthly values of the dependent and explanatory variables. Since pricing variation occurs in three-hour units of time each day, that breakdown of time is used to determine the unit of analysis for our monthly investigation. So, for example, the first pricing period in the day is 9 a.m. to noon, the second is noon to 3 p.m., the third is from 3 p.m. to 6 p.m., and the fourth is from 6 p.m. to 9 p.m.. A fifth period, 9 p.m. to 9 a.m., is used for the entire nighttime period, since no additional variation in price took place within those hours.

All of the exploratory analysis reported here is based on weekday consumption, Monday through Friday.

## III. EXPLORATORY ANALYSIS

Our first concern in examining the response to time-of-day pricing is to find robust and efficient estimates. Rather than selecting a specification based on a predetermined structural equation such as the translog or generalized Leontief system, we have allowed the data to suggest the appropriate specification. Transformations of the dependent variable were selected to yield a nearly normal error structure in order to produce well-behaved t and F statistics for hypothesis testing and robust coefficient estimates. Transformations of the independent variables were selected to reduce heteroskedasticity. Finally, we examined the data to detect outliers that might have unduly influenced the results.

The general form of our problem is to estimate the demand equations for the five time periods in the form

$$\gamma(kwh_{it}) = f_i\ (P_1, \ldots, P_5, X_1, \ldots, X_n, W_{it}) \qquad (1)$$

where i = 1, ..., 5; n = 1, ..., 100+; and $\gamma(kwh_i)$ is a transformation of the electricity use in time period i of month t. The P's are prices, the X's are other explanatory variables, and W is weather in month t. There are 18 months of observations and several possibly significant interactions (e.g., between price and appliance, between appliance and weather, and between all three variables for several of the Xs).

The specification of the model was broken into parts for reasons of cost and computational efficiency. First, transformations of the dependent variables were accomplished with Box-Cox techniques applied to a sample of the household-months in the data set. Second, appropriate specification of the right-hand-side variables was accomplished sequentially from an analysis of variance (ANOVA) specification of experimental rates to a more parsimonious price specification that is consistent with the ANOVA results. Finally, we applied the outlier measure proposed by Cook to see if the estimates were unduly influenced by outlier observations.

TRANSFORMATION OF THE DEPENDENT VARIABLE

In analyzing the effects of time-of-day pricing, we used a two-parameter Box-Cox transformation of kwh use by time of day. (See Box and Cox (1962).) The dependent variables were specified as

$$\frac{(\text{kwh} + c)^{\lambda-1}}{\lambda}, \quad \lambda \neq 0.$$

The limit of this expression is log(kwh + c) as $\lambda$ goes to 0, where "log" is the natural logarithm throughout this discussion. We adopted the logarithmic specification because the maximum likelihood estimate of $\lambda$ was insignificantly different from zero. Values of c were chosen in the middle of the region for which the log likelihood was maximized; the log likelihood function was very flat in the regions of c used here, and we selected 0.1. These values of c also have the property that retransforming the variable to raw kwh provides a relatively close estimate of the mean of the raw, untransformed data.[1]

The Box-Cox transformations selected have three important properties for our analysis. First, the transformed variable is approximately normally distributed; thus the resulting residuals from the regression should be approximately normal, leading to well-behaved t, F, and $\chi^2$ statistics. Second, the regression coefficients will be less affected by the unusually large values of consumption in the right tail of the distribution with this logarithmic transformation than would be true in running regressions on untransformed data. Electricity consumption in this sample is highly skewed, with a very thick or heavy right tail. Finally, the use of a logarithmic transformation leads to more efficient estimates of mean consumption because the coefficient of variation is smaller than with untransformed kwh.

The effect of the transformation can be seen in Fig. 1. Figure 1A is a normal plot for kwh consumption from 9 a.m. to noon, where the

[1] If the variable is not exactly log normal, there is an approximation error in using the formula $\exp(x\hat{\beta} + \hat{\sigma}^2/2)$ to estimate the mean of the untransformed kwh.

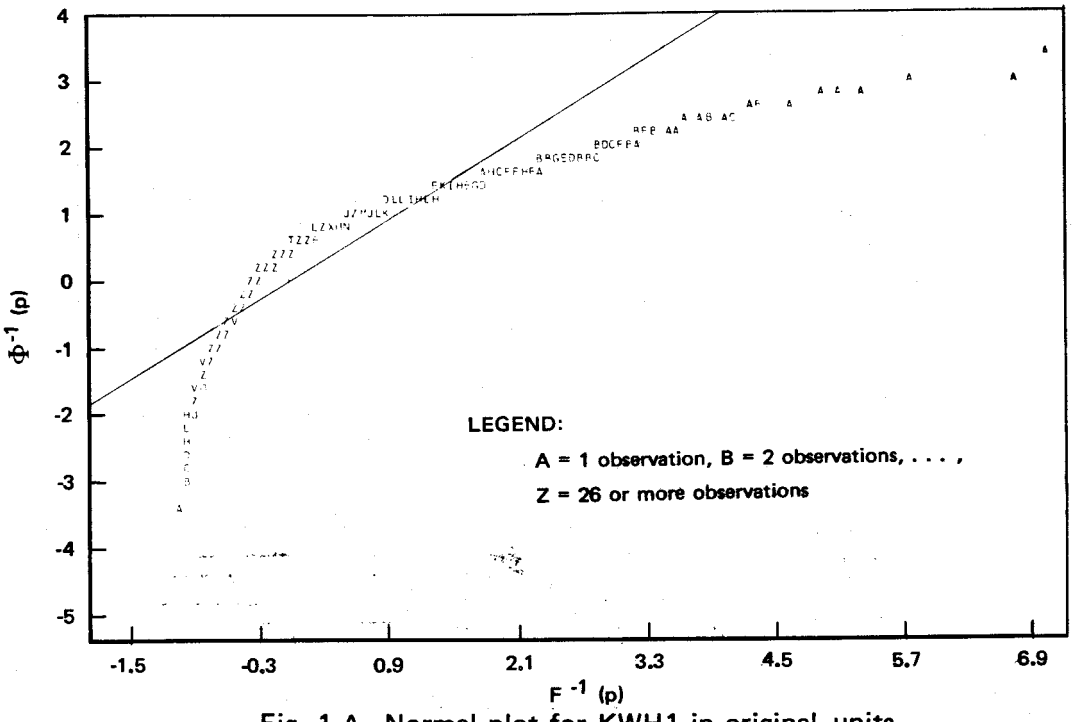


Fig. 1.A—Normal plot for KWH1 in original units
(9AM—noon, weekdays)

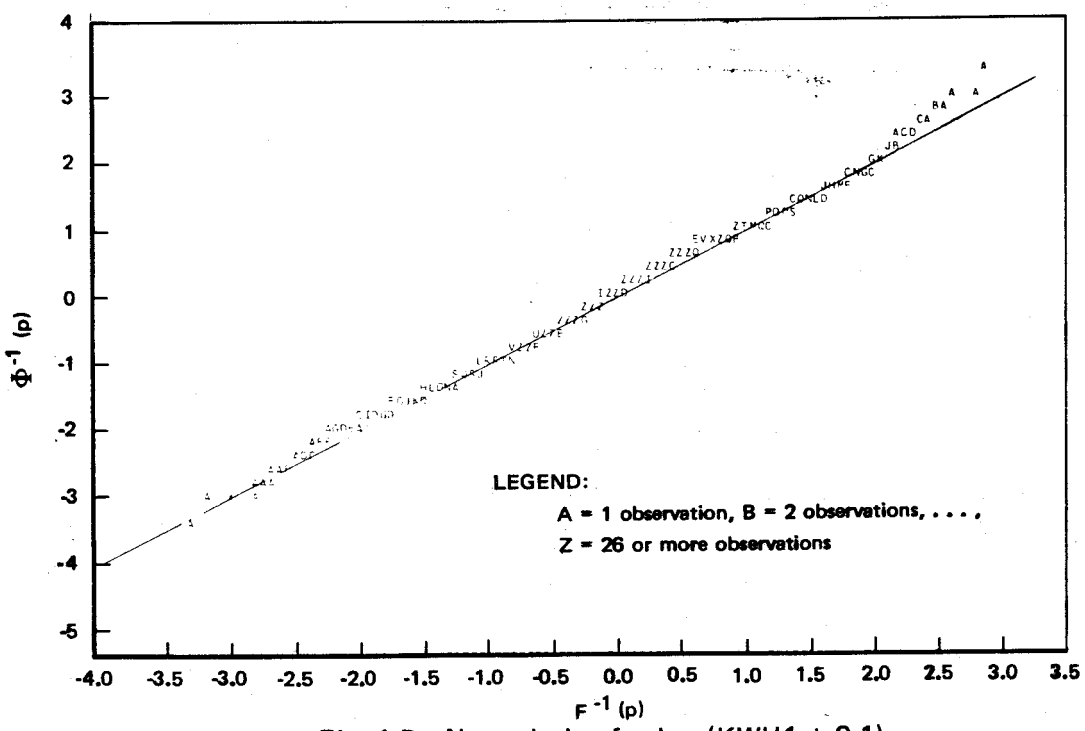


Fig. 1.B—Normal plot for log (KWH1 + 0.1)
(9AM—noon, weekdays)

horizontal axis has the inverse of the empirical cumulative distribution function (CDF), $F^{-1}(p)$ and the vertical axis gives values of the inverse of the theoretical CDF, $\phi^{-1}(p)$. When a variable is normally and symmetrically distributed, the data points lie on the straight line shown in Fig. 1A. Figure 1B is a normal plot for the transformed consumption in this time period. As Fig. 1B indicates, the transformations create symmetric and nearly normal dependent variables, ensuring that regression results will be robust.

TRANSFORMATIONS OF THE EXPLANATORY VARIABLES

With the exception of our dichotomous variables, almost all of the household variables used to explain an electricity consumption have a few observations with very large values. For example, there are several experimental households with incomes well in excess of $50,000 per year and home values in excess of $200,000, although the median values for these variables in our sample are $18,200 and $43,000 (in 1976 values). To reduce heteroskedasticity and undue influence of the small number of very large values, all explanatory variables (except the dichotomous ones) were transformed. We found that a logarithmic transformation of the independent variables produced fairly homoskedastic residuals when plotted against the independent variables.

To reduce estimating costs for this analysis, we averaged the monthly data on each household into a single average observation. This avoided the problem of addressing simultaneously the issues of deterministic model specification and serial correlation. The average observation used is

$$\bar{y}_{tim} = x'_{1tim}\beta_1 + \bar{x}'_{2tim}\beta_2 + \bar{\varepsilon}_{tim} \qquad (4)$$

where t = time-of-day period,

i = household,

m = month,

$y_{tim}$ = consumption in period t by household i in month m,

$x_{1ti}$ = column vector of household characteristics that do not vary by month,

$x_{2tim}$ = column vector of characteristics that vary by month,

$\varepsilon_{tim}$ = error term,

and where a bar over a variable indicates that it has been averaged over $M_i$ observations for household i.

This averaging process leads to a relatively small loss in precision with respect to variables that do not change from month to month--for example, household composition and appliance ownership. The chief loss in precision occurs with respect to weather-related variables, which vary monthly. Since our primary concern in the exploratory data analysis was to model the effects of price and appliances, none of which vary over time in the data available, this loss in precision is not major.[2]

Nonprice Variables

The in-person household survey provided information on more than 100 nonprice variables that can be used to explain electricity use and the response to time-of-day prices. Although the large amount of data in the Los Angeles experiment would permit including a very large number of explanatory variables, we prefer as parsimonious a specification as possible to facilitate interpretation of the results while still achieving a good fit of the data and accurate predictions.

---

[2] If the error structure in our data is of a variance components variety, with a correlation coefficient $\rho$, we can estimate the loss of precision for the price variables from the fact that the population weight $w_i$ for time-invariant variables is approximately $w_i = (1 - \rho + M_i\rho)/M_i$, where $M_i$ is the number of observations for the ith household. For typical values such as M = 15 and $\rho$ = 0.80, the standard error of the price coefficient is increased by 11 percent when using an average observation rather than the full data.

Table 2 lists the explanatory variables used in our analysis (including price variables). The important nonappliance variables are income (combined wage and nonwage), housing value, monthly rent, family size, number of rooms, and indicator variables for use of air conditioner during weekdays, and for times of the week when an adult is normally at home. A total of 16 appliance variables are available for use in the Los Angeles data. On the basis of other analysis,[3] we have found that interacting heating and air conditioning appliances with heating and cooling degree hours, respectively, produced reliable predictions of monthly electricity use.

We used logarithmic transformations of the explanatory variables, other than the indicator variables, in order to minimize heteroskedasticity. When an untransformed explanatory variable was used, there was a strong tendency to overfit the few large values for the variable. For example, Fig. 2A shows the plot of residuals to our estimating equation versus ROOMS when ROOMS is the explanatory variable. Figure 2B is the plot of residuals versus log(ROOMS), when the specification includes the transformed explanatory variable. The latter plot indicates a far more homoskedastic plot than the former does.

To achieve greater parsimony in specification, we tried an aggregation of nonclimate-related appliance variables (i.e., all appliances other than air conditioning, heating, and swimming pool). In addition to parsimony, aggregation is appealing in a logarithmic specification. It seemed inappropriate to estimate the proportional changes in consumption as a linear function of individual appliance counts. Adding the effect of a clothes dryer to a base consumption that includes only a TV and a regrigerator should yield a larger proportional change than would adding the dryer to the consumption level of an all-electric home. Using appliance weights developed in a Los Angeles load study to reflect average monthly use of the appliances, we aggregated 12 appliances into a single variable, APPL. That is,

---

[3] Acton, Mitchell, and Sohlberg (1980).

Table 2

VARIABLE DEFINITIONS

DEPENDENT VARIABLES

| | |
|---|---|
| K1 | log (kwh + 0.1) for 9 a.m.-noon weekdays |
| K2 | log (kwh + 0.1) for noon-3 p.m. weekdays |
| K3 | log (kwh + 0.1) for 3-6 p.m. weekdays |
| K4 | log (kwh + 0.1) for 6-9 p.m. weekdays |
| K5 | log (kwh + 0.1) for 9 p.m.-9 a.m. weekdays |

INDEPENDENT VARIABLES

Price Variables

| | |
|---|---|
| Tariff Dummies | Indicator variables for each of 19 rates |
| LP1 | log (price during 9 a.m.-noon weekdays) |
| LP2 | log (price during noon-3 p.m. weekdays) |
| LP3 | log (price during 3-6 p.m. weekdays) |
| LP4 | log (price during 6-9 p.m. weekdays) |
| LP5 | log (price during 9 p.m.-9 a.m. weekdays) |

Demographic Variables

| | |
|---|---|
| ACWD | runs air conditioner weekdays |
| FAM | log (family size) |
| HOMEWD | adult home weekdays |
| HOMEWE | adult home weekends |
| HOMEWN | adult home weeknights |
| INC | log (income) |
| OWNER | dummy variable for owner |
| RENT | log (rent + 1.0) |
| ROOMS | log (rooms) |
| VALUE | log (home value + 1.0) |

Weather Variables

| | |
|---|---|
| ACT | average daily cooling degree hours |
| AHT | average daily heating degree hours |

Appliance Variables

| | |
|---|---|
| AIRCDT | log (weighted sum of air conditioner X ACT + 1.0) |
| APPL | log (appliance aggregate); see fn. 4 for def. of weights. |
| BWTV | black-and-white TVs |
| COLORTV | color TVs |
| COMBSTV | combined electric/natural gas stoves |
| DISHWASHER | has a dishwasher |
| EDRYER | has an electric dryer |
| ESTV | electric stoves |
| FFRIG | main refrigerator is frost-free |
| FRIG | refrigerators |
| FREEZE | has a freezer |
| GDRYER | has a gas dryer |
| HEATT | log (electric space heat AHT + 1.0) |
| POOL | has a swimming pool (apartment house pools excluded) |
| SCSTV | main stove is self-cleaning |
| WASHER | has a clothes washer |

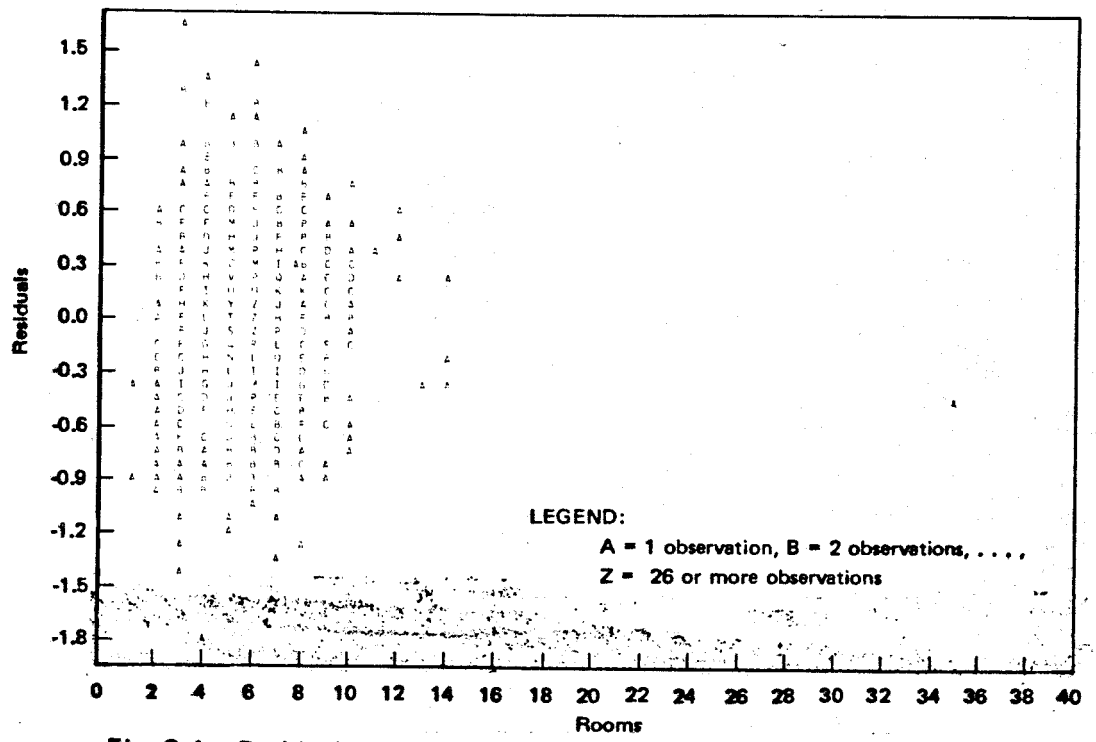


Fig. 2.A—Residual plot for log (KWH1 + 0.1) against untransformed rooms (9AM—noon, weekdays)

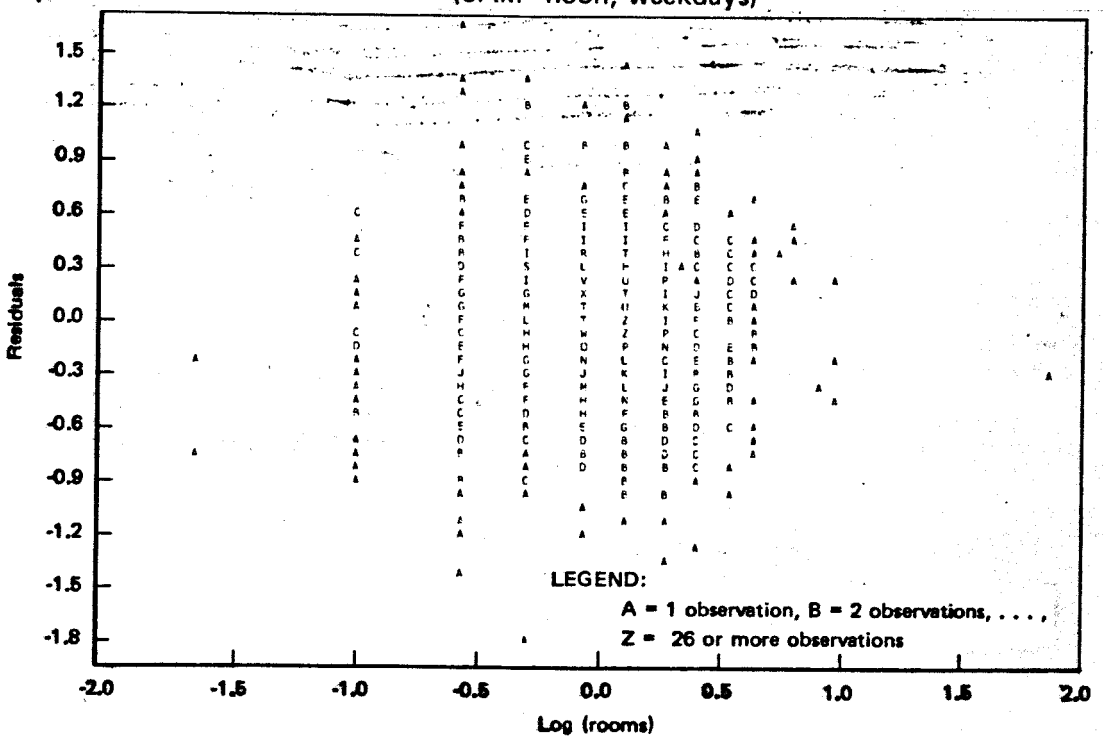


Fig. 2.B—Residual plot for log (KWH1 + 0.1) against log (rooms) (9AM—noon, weekdays)

$$APPL = \log \left[ \left( \sum_{i=1}^{12} \omega_i A_i \right) \right]$$

where $\omega_i$ is the weight and $A_i$ is the number of appliances of type i.[4] APPL was statistically significant in each of the time periods and, perhaps more important, the coefficients on remaining variables were insensitive to the change in specification. As a consequence, we simplified the estimation equation by using APPL in place of a separate listing of nonclimate-related appliances.

Transformation of Price

The specification of the price effects proceeded sequentially from an analysis of variance specification of experimental rates to a parametric representation of price effects. The ANOVA regressions contained a dummy variable for each of the experimental plans plus variables for appliances and other household characteristics. The dependent variable was average monthly values of log (kwh + 0.1) for each time-of-day period. The pattern of coefficients on the price dummies indicates that consumption is negatively related to the price in that period, but at a decreasing rate. Several transformations of the price variable could represent this decrease in proportional influence of price on consumption. One of those is the natural logarithm of price.

For each of the five time-period regressions, the analysis was performed by replacing the dummy variables with values of log price for all five pricing periods. We conducted an F test for changing the price specification from ANOVA to log (price). None of the calculated values are significant at 0.10 level, and the F values are about what would be expected at random.[5] Therefore, we accept the logarithmic

---

[4]The values of the weights used in this aggregation are frost-free refrigerator (106), manual defrost refrigerator (52), dishwasher (25), gas dryer (8), electric dryer (72), clothes washer (8), electric water heater (333), electric stove (87), combined electric/gas stove/oven (43), freezer (95), black and white TV (30), and color TV (30).

[5]The F values for the five equations are 0.79, 1.19, 0.91, 1.26, and 0.53, with degrees of freedom (13, 1113).

transformation as a convenient and familiar one that is consistent with the results of the analysis of variance.

In addition to calculating the F test, the residuals for each period of time were plotted against each of the price variables. The resulting plots were fairly homoskedastic. The residuals were also regressed on planned dummy variables to see if there were any systematic patterns of effect; none could be discerned. The combination of F tests and residual analysis suggested that the log (price) formulation fits well in spite of the heterogeneity of experimental plans: plans with time-invariant prices, plans with peak prices extending from three to twelve hours, and price levels ranging from 1¢ to 13¢ per kwh.

## ANALYSIS OF RESIDUALS AND OUTLIER OBSERVATIONS

We examined the residuals to determine if the model was specified correctly. The residuals from the regression on averaged data are symmetric and approximately normally distributed.[6] The plots of residuals versus major variables suggest that the residuals are fairly homoskedastic. There is some slight decrease in variance with increasing price. These plots suggest that the price function fits the data well without over-fitting particular experimental treatments. The residuals versus APPL plot suggests that the log of weighted-sums specification was well behaved.

Using the outlier measure proposed by Cook (1977), we examined the data to see if there were any other highly influential data points. Cook's distance is a measure of the extent to which dropping the $i^{th}$ observation affects the coefficient vector $\hat{\beta}$. If the $i^{th}$ observation is an outlier in both the dependent variable conditional on X and in the independent variables, then that observation can have a large effect on the estimated $\hat{\beta}$. The largest values were well below the 10-percent level suggested by Cook. However, a systematic examination indicated that a disproportionate percentage of the large Cook's distance values were associated with swimming pool ownership. Using only an indicator variable for the presence of a pool appears to be

---

[6] The distribution of residuals falls a little too much in the tails to accept the hypothesis of exact normality.

an incomplete specification. The log of the size of the pool's surface area has been added to reflect differences between large and small pools. When this adjustment is made, the pool owners no longer have disproportionately larger outlier measures.

The outlier examination did not suggest that the appliance aggregate (APPL) was misspecified. The presence of individual appliances among the outliers was about what would be expected for this sample.

## IV. ESTIMATION RESULTS FOR POOLED TIME-SERIES AND CROSS-SECTION DATA

Once we had developed a satisfactory initial specification of the model, we turned to the monthly data to obtain more efficient estimates. As long as the unobserved component in the error structure is independent of the observed independent variables, the average data provide consistent estimates. However, the loss of information on time-varying parameters (e.g., the weather-sensitive appliances), and an unequal number of observations, make ordinary least squares (OLS) estimates inefficient. This analysis proceeded in three steps.

First, because repeated time-series observations may contain serial correlation in the error term, we examined the data for intertemporal correlation. We concluded that the error structure could be approximated by a variance components model. We report estimates obtained using a random effects model, discussed below under "serial correlation."

Second, since the Los Angeles Rate Study is a voluntary social experiment, there is a possibility of selection bias. We examined the data for evidence of self-selection and concluded that an adjustment should be made to avoid misestimation of the price response.

Third, we examined the principal results of our preliminary analysis using the time-series data, paying particular attention to the main price effects and the appliance effects, and testing for price-appliance interactions that may have escaped detection in the analysis based on average data.

### SERIAL CORRELATION

Serial correlation can occur in the analysis of time-series data on individual households. Unmeasured tastes, appliance differences, and habits are likely to persist from one observation to another. As long as these unobserved components are not correlated with the independent variables and there are no lagged endogenous variables, ordinary least squares will provide consistent estimates of the

coefficient vector. However, least squares estimators will produce biased estimates of t and F statistics. For example, the calculated t statistics for time-invariant variables will overstate the statistical precision of the variables' effects.

An exploratory analysis of OLS residuals from the monthly data suggested that the household's error structure was similar to that exhibited by a variance components model of the form

$$y_{jt} = x'_{jt}\,\beta + \mu_j + \varepsilon_{jt} \qquad \begin{matrix} j = 1, \ldots, n \\ t = 1, \ldots, T_j \end{matrix} \tag{6}$$

or

$$Y = X\beta + \mu + \varepsilon,$$

where j indexes households and t indexes time. That is, the error structure appears to have a permanent unmeasured household component $\mu_j$.

For the monthly data, we have used a random effects model to account for serial correlation. In that model, the two error components $\mu$ and $\varepsilon$ are assumed to be independent of one another and uncorrelated with the observed independent variables. The household component $\mu$ is assumed to be $N(0, \sigma_\mu^2)$ and to be independently and identically distributed across households. The individual component is assumed to be $N(0, \sigma_\varepsilon^2)$ and to be independently and identically distributed across observations. The variance of the error is

$$V(\mu + \varepsilon) = (\sigma_\mu^2 + \sigma_\varepsilon^2) \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_n \end{pmatrix} \tag{7a}$$

where

$$A_j = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \vdots & \ddots & & \vdots \\ \rho & & \ddots & \rho \\ \rho & \cdots & & 1 \end{pmatrix} \qquad (T_i \times T_i) \tag{7b}$$

$$\rho = \sigma_\mu^2/(\sigma_\mu^2 + \sigma_\varepsilon^2) \tag{7c}$$

If no correlation exists, $\rho = 0$ and the OLS assumptions hold.

The random effects estimator uses between- and within-household accumulations of the variables. As a straightforward extension of Maddala's (1971) work, one can show that the GLS estimator for data with unequal numbers of observations is

$$\hat{\beta}_{GLS} = \left[ \sum_{\tau=1}^{T} W_{XX\tau} + \theta_\tau B_{XX\tau} \right]^{-1} \left[ \sum_{\tau=1}^{T} W_{XY\tau} + \theta B_{XY\tau} \right] \qquad (8a)$$

$$V(\hat{\beta}_{GLS}) = \sigma^2 \left[ \sum_{\tau=1}^{T} W_{XX\tau} + \theta_\tau B_{XX\tau} \right]^{-1} \qquad (8b)$$

where $$\theta_\tau = \frac{1-\rho}{1-\rho+\rho\tau} \qquad (8c)$$

$$B_{XX\tau} = \frac{1}{N_\tau} \sum_{i=1}^{N_\tau} X_i' ee' X_i$$

$$W_{XX\tau} = \sum_{i=1}^{N_\tau} X_i' X_i - B_{XX\tau} \qquad (8e)$$

$N_\tau$ = number of households with exactly $\tau$ observations, and

e = a column vector of $\tau$ 1's .

$B_{XY}$, $W_{XY}$ are similarly defined.

Because the value of $\rho$ is unknown, it must be estimated. For the results below, we used maximum likelihood to estimate the coefficients $\beta$ and correlation $\rho$.[1]

The intraclass correlation coefficients ($\rho$) are presented in Table 3. The results include estimates from analysis based on

[1]We are indebted to Dan Relles of Rand for providing us computationally efficient software for the random effects model with unequal observations per household.

Table 3

ESTIMATES OF INTRACLASS CORRELATION COEFFICIENTS

| | Uncorrected for Self-Selection | | Corrected for Self-Selection | |
|---|---|---|---|---|
| Time Period | ρ | Sum of Weights[a] | ρ | Sum of Weights[a] |
| 9 a.m.-noon | .7145 | 1550 | .7046 | 1569 |
| Noon-3 p.m. | .6796 | 1620 | .6778 | 1623 |
| 3-6 p.m. | .6146 | 1767 | .6117 | 1775 |
| 6-9 p.m. | .5958 | 1816 | .5948 | 1818 |
| 9 p.m.-9 a.m. | .7502 | 1485 | .7496 | 1486 |

[a]Sum of weights = number of effectively independent observations for variables that are time-invariant and are uncorrelated with time-varying variables.

uncorrected data and for data corrected for self-selection--discussed in the next subsection. The effect of such high correlation is to reduce the sample size from 15,062 observations to 1500 to 1800 effectively independent observations when analyzing time-invariant variables such as experimental price.

TESTING FOR SELF-SELECTION BIAS

Households contacted for the Los Angeles Study had the right to refuse participation and remain on the conventional declining block rate, but acceptance was very high (more than 90 percent).[2] Lump-sum compensation was made to the approximately one-third of households who faced plans that would have made them worse off at preexperimental levels of use.[3] Even with such compensation available, some households with relatively inelastic demand may have refused the less advantageous plans (those with high peak prices or long peak periods) if they perceived a possible financial loss. Such self-selection, if it took place, could lead to overestimates of price elasticities in uncorrected estimation equations.

We tested for selection bias using techniques developed by Heckman (1976) and Amemiya (1973), which have been widely used in the analysis of labor participation.[4] For simplicity in describing the procedure, consider a two-equation system involving an enrollment acceptance decision (where the household joins the experiment if $A > 0$) and a consumption amount decision K:

$$A = X\beta_A + \nu_A \tag{9a}$$

$$K = X\beta_K + \nu_K \tag{9b}$$

where the errors may be correlated across equations but not across households. The errors are distributed as a bivariate normal. The selection bias occurs if the enrollment decision is correlated with consumption error and the observed Xs.

---

[2] See Chow and Mitchell (1979).

[3] See Manning, Mitchell, and Acton (1979) for a description of the compensation scheme, which was based on one year's preexperimental consumption. Households were offered only one predetermined experimental plan. The compensation was determined as the difference between the electricity bill under the conventional rate and the bill at the experimental rate at an unchanged level and pattern of consumption.

[4] More recently, Aigner and Hausman (1980) have applied similar techniques to the data from the Arizona electricity experiment.

To correct for selection, we use the inverse Mill's ratio,

$$\lambda_i = \frac{f\left(\frac{x_i' \beta_A}{\sigma_{AA}}\right)}{1 - F\left(\frac{x_i' \beta_A}{\sigma_{AA}}\right)} \tag{10}$$

where f and F are the normal probability and cumulative distributions for the household's preexperimental consumption. A predicted value of $\lambda$ is used to correct the consumption equation by adding it to the independent variables:

$$K = X\beta_K + \hat{\lambda}\alpha + \varepsilon \tag{11}$$

As Amemiya (1973) has shown, the resulting two-stage estimates are consistent. In order for Eq. 11 to be identified, the explanatory variables for the instrument $\lambda$ must not be identical to the Xs in Eq. 11.

In estimating Eq. 9a, we allowed for a large number of plan-related and household-specific variables. In addition to price levels, we had several indicator variables for plan type (flat, seasonal, TOD, and 5-day or 7-day peak prices), amount of participation payment, ethnicity and locational variables, appliances, income, and preexperimental level of use. Since preliminary analysis suggested that log of preexperimental kwh and income entered nonlinearly, we entered both in quadratic form and allowed an interaction of income and log kwh.

Results of this analysis indicated a relatively small plan-related selection bias. Higher prices carried greater probability of refusal, but only one TOD coefficient was statistically significant at the 0.05 level. Households requiring participation payment were more likely to refuse. Households with electric space heating and with air conditioning were slightly more likely to refuse. Finally, the probability of acceptance increased with income (at an increasing rate). Thus,

refusals were disproportionately poor; in particular, low-income households with higher preexperimental consumption were more likely to refuse to join the study.

When $\lambda$ is entered in the estimation equations reported below, it is highly statistically significant and negative, indicating the importance of making the self-selection correction. Using the $\lambda$ correction reduces some of the own-price elasticities and causes several cross-price elasticities to have the correct, positive, sign.

## TWO MODELS OF TOD DEMAND

We can now examine the principal time-series cross-sectional results, using the corrections for selection bias and serial correlation just discussed. We report two models: first, a relatively simple main effects model in which all variables, including own and cross prices, enter additively in logs; second, a model with interactions between prices and the major appliances. Since it is relatively difficult to see the complete effects of a factor (such as air conditioning) in the highly interactive model, we concentrate our discussion on the contribution of each main factor in the simple model.

### The Main Effects of Price

The simple model takes the general form

$$\log (kwh_i + 0.1) = a_{i0} + \alpha_i \lambda + \sum_{j=1}^{5} a_{ij} \log p_{ij} + \sum_{j=1}^{4} b_{ij} DZ + \sum_{j=1}^{4} d_j \log Z_j + \sum_j e_{ij} X_j \qquad (12)$$

where the variables are defined as in Table 4. Table 5 presents the own and cross price effects for weekday consumption. Own-price effects

Table 4

DEFINITIONS OF VARIABLES IN ESTIMATED DEMAND EQUATIONS

DEPENDENT VARIABLE

$K_i$ = daily kwh in period i.

INDEPENDENT VARIABLES

$P_j$ = price in period j.

$Z_1$ = 0 if no air conditioning; otherwise = log(AC · CDH),

where AC = $\sum_i \omega_i A_i$ for four air conditioning appliances,[a] and

CDH = cooling degree hours measured as deviations from 65° F.

$Z_2$ = 0 if no HEAT; otherwise = log ROOMS · log(HEAT · HDH),

where HEAT = electric space heating, and

HDH = heating degree hours measured from 65° F.

$Z_3$ = 0 if no swimming pool; otherwise = log (length times width of pool).

$Z_4$ = log $\Sigma\omega_i A_i$ for all other appliances,

where $A_i$ = count of appliance i in the household,

$\omega_i$ = average kwh for appliance j,[a] and

$\lambda$ = the inverse Mill's ratio for selection bias correction.

---

[a] Values of $\omega$ are given in footnote 4, Sec. III.

Table 5

ESTIMATED OWN- AND CROSS-PRICE ELASTICITIES

(|t| in parentheses)

| Period | Price | | | | |
|---|---|---|---|---|---|
| | 9 a.m.-Noon | Noon-3 p.m. | 3-6 p.m. | 6-9 p.m. | 9 p.m.-9 a.m. |
| 9 a.m.-noon | -.1047<br>(5.93) | -.0264<br>(1.07) | +.0478<br>(2.11) | +.0750<br>(2.66) | +.0804<br>(2.39) |
| Noon-3 p.m. | +.0043<br>(0.23) | -.1020<br>(3.95) | +.0128<br>(0.54) | +.0547<br>(1.85) | +.0278<br>(0.79) |
| 3-6 p.m. | +.0149<br>(0.88) | +.0041<br>(0.17) | -.0774<br>(3.57) | +.0365<br>(1.35) | +.0297<br>(0.63) |
| 6-9 p.m. | +.0109<br>(0.77) | -.0005<br>(0.02) | -.0153<br>(0.84) | -.0509<br>(2.26) | -.0099<br>(0.36) |
| 9 p.m.-9 a.m. | -.0065<br>(0.42) | +.0128<br>(0.59) | +.0164<br>(0.82) | +.0191<br>(0.77) | -.0581<br>(1.96) |

are relatively small, but highly statistically significant. They average about -.10 in the morning and early afternoon, fall to -.08 in the late afternoon and fall even lower in the evening and overnight. Almost all of the cross-price effects have the expected positive sign but, with the exception of the equation for 9 a.m. to noon, cross-price coefficients are not statistically significant.

Table 6 shows the effects of making the self-selection correction. The coefficient $\lambda$ is negative and very significant statistically in each of the five equations. The own-price coefficients are essentially unchanged in the morning and early afternoon but are notably smaller in the other three time periods when the selection adjustment is made.[5]

[5]When the self-selection correction is omitted, the cross-price effects also change, with only one coefficient statistically different from 0 and many taking on the incorrect negative sign.

Table 6

VALUES OF SELECTION CORRECTION PARAMETER, $\lambda$, AND EFFECTS ON ESTIMATED OWN-PRICE COEFFICIENTS

| | Own-Price Coefficients | | |
|---|---|---|---|
| Time Period | Estimated Without $\lambda$ | Estimated With $\lambda$ | Value of $\lambda$ |
| 9 a.m.-noon | -.1001 (5.57) | -.1047 (5.93) | -.5310 (2.94) |
| Noon-3 p.m. | -.1109 (4.30) | -.1020 (3.95) | -.6025 (3.18) |
| 3 p.m.-6 p.m. | -.1125 (5.71) | -.0774 (3.57) | -.6595 (3.81) |
| 6 p.m.-9 p.m. | -.0851 (4.94) | -.0509 (2.26) | -.3387 (2.34) |
| 9 p.m.-9 a.m. | -.0808 (2.90) | -.0581 (1.96) | -.3489 (2.19) |

This finding is consistent with the predicted effects of a selection bias, and suggests that it is important to make the correction even in an experiment that had an enrollment acceptance rate well in excess of 90 percent.

Appliances and Other Variables

Appliances holdings and weather strongly affect the quantity of electricity used in each time period. Table 7 reports the effects of major appliance variables. Air conditioning and electric space heating are entered as a (0, 1) dummy variable and as a count of appliances multiplied by daily cooling degree hours or heating degree hours averaged over the month. Each of these variables is statistically significant and exhibits the expected daily pattern of use. Air conditioning is proportionally more responsive to changes in temperature during the middle of the day than before noon or after 6 p.m. Conversely, electric heating is most responsive to a given value of heating degree hours in the overnight period.

Table 7

RESPONSES TO SELECTED APPLIANCE VARIABLES

(|t| in parentheses)

| | Air Conditioning | | Electric Heat | | | Swimming Pool | |
|---|---|---|---|---|---|---|---|
| Time Period | ARCD>0 | ARCDT | HEAT>0 | HEATT | APPL | POOL>0 | SIZE |
| 9 a.m.-noon | -.5864 | .0776 | -.2755 | .0780 | .5554 | -1.0475 | .2604 |
| | (13.17) | (21.50) | (4.01) | (10.15) | (17.97) | (2.01) | (3.19) |
| Noon-3 p.m. | -1.3770 | .1821 | -.2516 | .0649 | .5127 | -1.0153 | .2423 |
| | (28.74) | (45.33) | (3.47) | (7.58) | (15.81) | (1.86) | (2.83) |
| 3-6 p.m. | -1.2680 | .1711 | -.2637 | .0857 | .5171 | -0.7712 | .1741 |
| | (27.25) | (40.91) | (3.92) | (9.58) | (17.49) | (1.54) | (2.23) |
| 6-9 p.m. | -0.5483 | .0784 | -.2514 | .0939 | .4215 | -0.7601 | .1499 |
| | (13.84) | (21.67) | (4.44) | (12.11) | (17.02) | (1.82) | (2.28) |
| 9 p.m.-9 a.m. | -0.1526 | .0192 | -.2650 | .1298 | .4712 | -0.8928 | .1846 |
| | (4.04) | (6.70) | (4.41) | (21.33) | (17.26) | (1.94) | (2.56) |

The appliance aggregate APPL for all other appliances also has a highly statistically significant positive coefficient in each equation; the coefficients range from .42 to .56. That is, all other things equal, a 10-percent increase in kwh-weighted stock of appliances will lead to a 4- to 5-percent increase in utilization. Since the weights do not vary across time periods, there is some error in variable that biases the coefficient toward zero. For example, electric stoves probably are most important in TOD response around mealtime, but our appliance aggregate gives the variable equal weight throughout the day.

The swimming pool variables also have a daily pattern of use, exhibiting their lowest impact in the afternoon and evening, the time of greatest pool use. Apparently, homeowners run the pumps and filters when the pool is less likely to be in use.

Income has a very small estimated elasticity and is often statistically insignificant. In the demand for electricity, income

is expected to have its major effects through the purchase of appliances. Furthermore, it will be highly correlated with housing value or monthly rent. When the selection correction $\lambda$ is used, income is also correlated with that variable. The remaining variables, such as housing value, family size, and number of rooms, exert the expected effect in period-by-period consumption.[6]

A Model With Price Interactions

The results just discussed for the simple model assume identical price response for all households. We might expect, however, that some households would have appliances whose use could be rescheduled. Thus a household's price response may be a function of appliance holdings. We examined this hypothesis with a model that contains interactions between each of the five TOD prices and each of the appliance stock variables in each of the five equations. The general form of the equation is

$$\log (kwh_i + .1) = a_{i0} + a_i\lambda + \sum_{k=1}^{5} a_{ik} \log p_k + \sum_{k=1}^{5}\sum_{j=1}^{4} b_{ikj} \log p_k DZ_j$$

$$+ \sum_{k=1}^{5}\sum_{j=1}^{4} c_{ikj} \log p_i \log Z_i + \sum_j d_{ij} \log Z_j$$

$$+ \sum_j e_{ij} X_j \qquad (13)$$

In general, the t statistics on individual coefficients are very small and often considerably less than 1. However, when we perform an F test for omitting a given set of interactions from a particular equation, we reject the hypothesis that the set of coefficients is equal to 0 in most cases.

---

[6]The complete set of coefficients from alternative specifications is available from the authors upon request.

Table 8

F TESTS FOR NO PRICE/APPLIANCE INTERACTIONS

| Time Period | Air Conditioning[a] | Electric Space Heating[b] | Other Appliances[c] | Swimming Pool[d] | All |
|---|---|---|---|---|---|
| 9 a.m.-noon | 1.84 | 2.95 | 1.60 | 5.41 | 3.21 |
| Noon-3 p.m. | 4.52 | 2.58 | 3.18 | 2.84 | 3.61 |
| 3-6 p.m. | 7.98 | 0.97 | 2.25 | 2.46 | 4.02 |
| 6-9 p.m. | 6.25 | 0.65 | 1.01 | 2.08 | 2.98 |
| 9 p.m.-9 a.m. | 0.77 | 11.97 | 2.60 | 1.36 | 4.45 |
| df | 10, 15002 | 10, 15002 | 5, 15002 | 10, 15002 | 35, 15002 |

Table 9

F TESTS FOR NO CROSS-PRICE INTERACTIONS WITH APPLIANCES

| Time Period | Air Conditioning[a] | Electric Space Heating[b] | Other Appliances[c] | Swimming Pool[d] | All |
|---|---|---|---|---|---|
| 9 a.m.-noon | 2.04 | 2.11 | 1.82 | 2.30 | 2.21 |
| Noon-3 p.m. | 2.12 | 2.13 | 3.65 | 0.54 | 1.83 |
| 3-6 p.m. | 4.08 | 1.04 | 2.29 | 1.05 | 2.24 |
| 6-9 p.m. | 3.72 | 0.85 | 1.18 | 0.89 | 1.88 |
| 9 p.m.-9 a.m. | 0.97 | 14.77 | 3.00 | 0.84 | 5.18 |
| df | 8, 14998 | 8, 14988 | 4, 14988 | 8, 14988 | 28, 14998 |

[a] Five prices X AIRCD>0 and AIRCDT.

[b] Five prices X HEAT>0 and HEATT.

[c] Five prices X APPL.

[d] Five prices X POOL and SIZE.

The effects on predicted price elasticity from a highly interactive model can be illustrated by calculating the value at mean levels of explanatory variables for different levels of monthly use. In a specification of the model where only own prices were interacted with the DZ's and Z's, the own price elasticity for a household in group k is given as

$$\eta^{own}_{ik} = a_{ii} + \sum_{j=1}^{4} b_{ij} \overline{DZ}_{jk} + \sum_{j=1}^{4} c_{ij} \overline{Z}_{jk} \qquad (14)$$

Table 10, adopted from Acton and Mitchell (1980), illustrates the difference in these elasticities as an average household moves from a low level of monthly use, characterized by minimal average holding of appliances, to high level of monthly use, characterized by high average values of APPL, air conditioning, electric space heating, and a large swimming pool. These dramatic increases in own-price elasticities represent almost a five-fold increase during some daylight hours, and a much smaller, but still large, differential in other periods. The effect of adding a swimming pool to households whose other characteristics remain the same is illustrated at mean values for households in the 1001-1100 kwh/month stratum.

These findings present us with an analytic and statistical choice. The richer specification with a great number of interactions is appealing because it allows price elasticities to vary with appliance holdings, and we find statistical support for those interactions. However, the interactions model contains a large number of variables that are fairly collinear. Thus the individual coefficients are very imprecisely estimated. Furthermore, the model with interactions is cumbersome to use in many applications. On the other hand, the simpler specification with no interaction is convenient but averages out a great deal of the variation across households in responsiveness to time-of-day pricing.

Table 10

OWN-PRICE ELASTICITIES FOR WEEKDAY ELECTRICITY USE BY LEVEL OF USE

| Household Consumption (kwh/mo) | Time Period 9 a.m.- Noon | Noon- 3 p.m. | 3 p.m.- 6 p.m. | 6 p.m.- 9 p.m. | 9 p.m.- 9 a.m. |
|---|---|---|---|---|---|
| **All Households** | | | | | |
| 0-200 | -.088 | -.008 | -.001 | -.012 | -.056 |
| 201-400 | -.075 | -.051 | -.032 | -.018 | -.038 |
| 401-500 | -.074 | -.066 | -.049 | -.031 | -.039 |
| 501-600 | -.089 | -.105 | -.069 | -.040 | -.035 |
| 601-700 | -.093 | -.128 | -.090 | -.048 | -.039 |
| 701-800 | -.125 | -.166 | -.112 | -.073 | -.051 |
| 801-900 | -.163 | -.222 | -.151 | -.091 | -.074 |
| 901-1000 | -.160 | -.212 | -.136 | -.092 | -.060 |
| 1001-1100 | -.197 | -.260 | -.167 | -.117 | -.082 |
| 1101-1200 | -.246 | -.313 | -.202 | -.150 | -.089 |
| 1201-1300 | -.231 | -.295 | -.202 | -.139 | -.077 |
| 1301-1400 | -.259 | -.341 | -.225 | -.162 | -.116 |
| 1401-1500 | -.265 | -.349 | -.250 | -.167 | -.083 |
| 1501-2000 | -.274 | -.368 | -.259 | -.174 | -.073 |
| 2001-2500 | -.309 | -.417 | -.400 | -.210 | -.062 |
| 2501+ | -.259 | -.405 | -.338 | -.180 | -.035 |
| **Households with Pools[a]** | | | | | |
| 1001-1100 | -.328 | -.395 | -.235 | -.189 | -.135 |

[a]Calculated at mean value of all nonpool explanatory variables for one stratum of use. Mean pool size for households having pools is used in each stratum prediction.

## DIRECTIONS FOR FUTURE ANALYSIS

The models just presented yield estimates that are highly satisfactory by most criteria. They fit the data well, have consistent tests of statistical significance, and yield parameter estimates that make sense in light of a priori predictions and correspond well to other published results. Two extensions of this work would be particularly attractive.

First, it may be useful to have a finer breakdown of estimated appliance effects within time periods. In particular, when the aggregate appliance variable APPL is created, it would be more satisfactory to have weights that vary by time period. This would lead to more accurate within-period estimates of appliance-specific effects.

Second, it is desirable to integrate this "model free" set of statistical demand curves with the implications of economic theory, including perhaps some of the restrictions of system demand curves based on utility maximizing behavior. Since the essential statistical properties of the data have been established, we can move with greater confidence into more highly structured models whose estimating equations often contain transformations of explanatory and dependent variables that are very similar to those tested. For example, the translog utility system explains consumption shares as a function of log transformations of explanatory variables. We now have some confidence that such transformations produce reasonable estimates with this data set and that the coefficients are not unduly influenced by outlier observations. Using such systems of demand equations will permit us to impose a greater structure on the set of coefficients and to derive more powerful implications about the partial and full price elasticities of demand for elasticity in different periods of time.

## V. CONCLUSION

In this analysis we have sought statistically reliable estimates of response to time-of-day (TOD) pricing. To do so, we have concentrated on statistical issues rather than solely economic ones. We sought transformations of the dependent and explanatory variables that would lead to reliable tests of statistical significance and efficient coefficient estimates. The result was to select a double logarithmic specification in which log of kwh is made a function of such variables as log of price and log of appliance holdings. It is very unlikely that, a priori, we would have selected such a specification, which implies constant elasticity of demand. Nevertheless, the results are highly statistically significant and the model leads to good predictions.

Our estimates also correspond well with the price elasticities estimated by other researchers in more highly structured models applied to other experimental data sets. For example, Aigner and Hausman (1980), Caves and Christensen (1980), and Taylor (1977) report comparable own-price elasticities from the Wisconsin and Arizona experiments.[1] The principal defects also correspond well to the results reported by Mitchell and Acton (1980), who examined a larger number of months of data from the Los Angeles experiment in a hybrid demand system of equations.

The magnitudes of price elasticities from this exploratory analysis of the Los Angeles data are quite small. On average, own-price elasticities are -.05 to -.10, suggesting that at least in the short run there may be relatively little welfare gain from introducing TOD pricing for most residential customers. Acton and Mitchell (1980) have used the results of this exploratory analysis to evaluate the effects of introducing TOD rates in the LA system. They find that, with an assumed

---

[1] Our values, however, are considerably smaller than those reported by Atkinson (1977, 1978), who used a translog system on the Arizona data.

metering cost of $150 per household, a TOD rate based on marginal costs in the Los Angeles system would produce a welfare gain of about $4.4 million per year at a cost of $17 million per year if applied to all of the one million residential customers. Thus, on the basis of efficiency and welfare criteria, TOD rates should not be applied universally in Los Angeles at these meter costs. But, given the elasticities in Table 10, which show that own-price elasticities depend on level of use and appliance holdings, selective application of TOD rates would produce a welfare gain in excess of metering costs. If TOD rates were applied only to households with average monthly consumption in excess of 1,100 kwh per month (less than 5 percent of Los Angeles's residential users, but accounting for 17 percent of consumption), then a net welfare gain of $3.3 million per year could result from metering costs of $1.4 million per year.

Obviously, these welfare comparisons will vary from one situation to another as levels of use, appliance holdings, weather, and the terms of the TOD rate are adapted to individual situations. The results demonstrate, however, that a flexible set of detailed parameter estimates such as those presented above can yield important information to assist evaluation of changes in electricity rates.

## REFERENCES

Acton, Jan Paul, and Bridger M. Mitchell, "Evaluating Time-of-Day Electricity Rates for Residential Customers," in B. M. Mitchell and P. R. Kleindorfer, eds., *Regulated Industries and Public Enterprise*, Lexington Books, Lexington, Mass., 1980, pp. 247-273.

Acton, Jan Paul, Bridger M. Mitchell, and Ragnhild Sohlberg, "Estimating Residential Electricity Demand Under Declining-Block Tariffs: An Econometric Study Using Micro-data," *Applied Economics*, Vol. 12, No. 2, June 1980, pp. 145-162.

Aigner, Dennis J., and Jerry A. Hausman, "Correcting for Truncation Bias in the Analysis of Experiments in Time-of-Day Pricing of Electricity," *The Bell Journal of Economics*, Vol. 11, No. 1, Spring 1980, pp. 131-142.

Aigner, Dennis J., and Dale J. Poirier, *Electricity Demand and Consumption by Time of Day: A Survey*, Palo Alto: Electric Power Research Institute, Report EA-1294, 1980.

Amemiya, Takeshi, "Regression Analysis When the Dependent Variable is Truncated Normal," *Econometrica*, Vol. 41, 1973, pp. 997-1016.

Atkinson, Scott E., "Responsiveness to Time-of-Day Electricity Pricing: First Empirical Results," in Anthony Lawrence, ed., *Forecasting and Modeling Time-of-Day and Seasonal Electricity Demand*, Palo Alto: Electric Power Research Institute, Report EA-578-SR, 1977.

-----, "A Comparative Analysis of Consumer Response to Time-of-Use Electricity Pricing: Arizona and Wisconsin," paper presented at the EPRI Workshop in Modeling and Analysis of Electricity Demand by Time of Day, San Diego, Calif., June 12-14, 1978.

Box, G. E. P., and D. R. Cox, "An Analysis of Transformations," *Statistical Society, Series B*, Vol. 26, 1964, pp. 211-243.

Caves, Douglas W., and Laurits R. Christensen, "Residential Substitution of Off-Peak for Peak Electricity Usage Under Time-of-Use Pricing," *The Energy Journal*, Vol. 1, No. 2, April 1980, pp. 85-142.

Chow, Winston, and Bridger M. Mitchell, *Sample Selection in the Los Angeles Electricity Rate Study*, The Rand Corporation, R-2430-DWP, Santa Monica, Calif., 1979.

Conlisk, John, and Harold Watts, "A Model for Optimizing Experimental Designs for Estimating Response Surfaces," *Proceedings of the Social Statistics Section,* American Statistical Association, 1969, pp. 150-156.

Cook, R. Dennis, "Detection of Influential Observation in Linear Regression," *Technometrics,* Vol. 19, February 1977, pp. 15-18.

Heckman, James J., "The Common Structure of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic and Social Measurement,* Vol. 5, A11 1976, pp. 475-492.

Kohler, Daniel F., *Essays in Applied Demand Analysis,* Ph.D. Dissertation, University of Michigan, 1980.

Lillard, Lee A., and Jan Paul Acton, *Seasonal Electricity Demand: A Variable Response Model,* The Rand Corporation, R-2425-DWP, Santa Monica, Calif., 1980.

Maddala, G. S., "The Use of Variance Components Models in Pooling Cross Section and Time Series Data," *Econometrica,* Vol. 39, March 1971, pp. 341-358.

Manning, Willard G., Bridger M. Mitchell, and Jan Paul Acton, "Design of the Los Angeles Peak-Load Pricing Experiment for Electricity, *Journal of Econometrics,* Vol. 11, No. 1, September 1979, pp. 131-194.

Mitchell, Bridger M., and Jan Paul Acton, *Electricity Consumption by Time-of-Use in Hybrid Demand System,* The Rand Corporation, R-2628-DWP, Santa Monica, Calif., 1980.

Morris, Carl, "A Finite Selection Model for Experimental Design of the Health Insurance Study," *Journal of Econometrics,* Vol. 11, No. 1, September 1979, pp. 43-62.

Taylor, Lester D., "On Modeling the Residential Demand for Electricity by Time of Day," in Anthony Lawrence, ed., *Forecasting and Modeling Time-of-Day and Seasonal Electricity Demand,* Palo Alto: Electric Power Research Institute, Palo Alto, Calif., 1977.