Volume Title: Evaluation of Econometric Models

Volume Author/Editor: Jan Kmenta and James B. Ramsey, eds.

Volume Publisher: Academic Press

Volume ISBN: 978-0-12-416550-2

Volume URL: http://www.nber.org/books/kmen80-1

Publication Date: 1980

Chapter Title: The Use of Exploratory Methods in Ecomonic Analysis: Analyzing Residential Energy Demand

Chapter Author: Lawrence S. Mayer

Chapter URL: http://www.nber.org/chapters/c11692

Chapter pages in book: (p. 15 - 45)

The Use of Exploratory Methods in Economic Analysis: Analyzing Residential Energy Demand

*LAWRENCE S. MAYER**

DEPARTMENT OF STATISTICS
PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY

1. Introduction

Exploratory data analysis is a set of methods designed to examine and describe the structure of data (McNeil, 1977; Mosteller & Tukey, 1977; Tukey, 1977). The exploratory approach, or mode, involves the use of descriptive techniques and the intuition of the analyst to investigate the structure of data. Unlike the confirmatory approach, or mode, which relies on classical statistical analysis and data to confirm and reformulate theory through the vehicles of parameter estimation and hypothesis testing, the exploratory mode uses data to suggest theory through the vehicle of informal data examination. It can be considered, in some sense, a prelude to classical statistical analysis. It is suggested that the exploratory mode is an important supplement to the confirmatory mode for the analysis of economic phenomena in areas, such as residential energy demand, in which the degree of prior belief in the available empirical theory or the quality and comprehensiveness of the available data may not be sufficient to allow adoption of the strong axiomatic structures associated with most classical statistical techniques. It is also suggested that the data analyst might consider combining exploratory data analysis and classical statistical analysis into a powerful analytic approach, or mode, which uses both formal statistical inference and

* Present address: Analysis Center, Wharton School, University of Pennsylvania, Philadelphia, Pa., 19104.

ISBN 0-12-416550-8

informal data exploration. To display the use of the exploratory mode, an exploratory analysis of data on residential energy demand is summarized. For contrast, a classical statistical analysis of residential demand is reviewed.

The data analyst, *au fait*, is capable of operating in several modes. The confirmatory mode, which relies on classical statistical methods, is appropriate for attacking problems for which the underlying substantive theory is well developed and has withstood repeated confrontations with data on similar problems. The theory yields the general form of the model of the relationships among the variables of interest. Strategy involves estimating the parameters of the model, testing certain hypotheses about the model, reformulating and re-estimating the model in light of the estimates and tests, and, ultimately, interpreting the parameter estimates and hypothesis tests.

For those problems in which the veracity of the available theory or models is not above suspicion, or no such theory or models are available, and for which the principal goal is to suggest hypotheses, models or theory, the exploratory mode, which relies on the methods of exploratory data analysis, is appropriate. The goal is to expose patterns in the data to obtain notions of the relationships among the variables of interest. These notions are used to suggest hypotheses, models or theories which may be candidates for future analysis in the confirmatory mode.

Finally, for those problems for which the available theory and models are worthy of testing but likely to be in need of significant alteration, the rough confirmatory mode, which combines classical statistical methods and exploratory methods, is most appropriate. Theory is used to formulate a model, parameters are estimated and hypotheses are tested, but emphasis is placed on both exploring the lack of fit of the model and interpreting the parameters and tests of the model. The object is to formulate a new, possibly radically different, theory and model which are proposed for further study but are not tested with the data at hand.

In the confirmatory mode the issue of statistical optimality is paramount. Theorems on the efficiency, unbiasedness, and asymptotic behavior of parameter estimates and the power, unbiasedness, and asymptotic behavior of test statistics are of primary importance. The philosophy of the mode is that the empirical theory and model adopted are sufficiently acceptable to eliminate, or at least postpone, the search for models radically different from the theory and model posited. The strategy is to use the theorems of mathematical statistics to estimate the parameters of the model and test hypotheses about the model, and then to use these estimates and tests to adjust the model. The parameters of the reformulated model are estimated and a final model is obtained. Rarely does the final model differ radically from the model first examined in terms of the variables included, functional form, or error specification.

For example, suppose that a well-tested empirical theory states that the log of demand for a particular good is linear, up to random error, in the log of real price, price is exogenously determined, and the assumptions of classical linear regression hold (independent errors, constant error variance, etc.), and suppose data on a random sample of individual consumers are available. Suppose that the analyst desires to estimate the model, test the hypothesis of no regression, and then test the hypothesis that the regression coefficient is equal to unity. It is clear that the confirmatory mode is appropriate. Using theorems of mathematical statistics the model is estimated and the hypotheses are tested. The theorems, hopefully, guarantee efficient estimators and the most powerful tests. Interpretation of the estimators and tests will probably be the major component of the interpretation of the results, although some attempt at model adjustment is permitted. It is important to remark that for problems such as these it would make no sense to forfeit the power and replicability of the confirmatory mode. The pivotal factor is the availability of a viable theory and suitable data.

The confirmatory mode is a method of analysis which is an integral part of the classical approach to science, an approach in which the testing of established empirical theory is paramount (Popper, 1968). Data are used to reject or revise theory. Scientific knowledge accumulates by disproving claims which arise from a posited theory.

In the exploratory mode visual examination of the data is emphasized. First, the distributions of the variables of interest are examined and loosely characterized in terms of notions such as location, variation, skewness, bimodality, extreme values, and shape. Then, the joint distributions of the variables or transformations of the variables are examined and patterns are identified and used to suggest hypotheses, models, and possibly theory. Such patterns may be in terms of conditional means, as are the patterns assumed in classical regression analysis, but may also be in terms of conditional variances, conditional skewness, conditional shape, conditional outliers, or other conditional characteristics which change systematically with regard to the exogenous variables. The philosophy is to embrace no theory or model other than that which is required to identify the problem of interest. The strategy is to use a variety of techniques which allow efficient dissection of the data. The end products are the display of the patterns found in the data, the residual of the data given the patterns, and possibly a tentative interpretation of the patterns. No inferences are made regarding any units of analysis other than those in the sample (Mayer, 1978a).

For example, suppose demand for a particular good is thought to be related to a variety of exogenous variables, including price, income, attitudes, demographic variables, and informational variables, but no theory is available to suggest either which factors are the most important or the shape of the relationship among the variables. Suppose that the analyst desires to examine the data relating the variables in order to suggest patterns worthy of testing in future studies. The exploratory mode is appropriate. The individual distributions and joint distributions of the variables are examined and patterns are noted. No inferences about a general population are made, no parameters are formally estimated, and no hypotheses are formally tested (although estimates and tests could be computed as exploratory guides). The goal is to separate the data into two components, the smooth (systematic) and the rough (residual). The rough displays no significant patterns but indicates variability. The smooth is suggested as a model and is used to formulate theory and hypotheses worthy of testing in additional, probably confirmatory, studies.

In the rough confirmatory mode a model is tentatively adopted and then tested using the methods of the confirmatory mode. Regardless of whether or not the model is rejected, the deviations from the model are examined using exploratory methods (e.g., see Welsch & Kuh, 1977; Welsch, 1977; Hoaglin & Welsch, 1978). The philosophy is to use the confirmatory mode to whatever degree possible but to use exploratory methods to obtain insights into the lack of fit of the model. The strategy is to adopt a model, estimate parameters, test the hypotheses, and then use exploratory methods to examine the residuals.

For example, suppose the analyst is willing to entertain, but not embrace, a model which states that the demand for a particular good is linear in price. The analyst may feel the model is good enough to deserve to be tested but likely to need significant modification. The rough confirmatory mode is appropriate. The model is estimated and hypotheses are tested. The residuals are explored. Interpretation of the results involves interpretation of the parameters and tests and interpretation of the patterns found in the residuals. The model may be radically reformulated after the examination of the residuals but is not reestimated by the same data. The rough confirmatory mode is like the confirmatory mode in that it begins with a model which is estimated and tested but like the exploratory mode in that the end product is new models, hypotheses, and theories which require testing on additional data.

Over the last decade many studies have used the confirmatory mode to develop models of residential energy demand (e.g., Anderson, 1972; Baughman & Joskow, 1975; Cohn, Hirst, & Jackson, 1977; Fisher & Kaysen, 1962; Griffin, 1974; Halvorsen, 1975; Hirst, Lin, & Cope, 1976; Mount, Chapman, & Tyrrell, 1973; Taylor, Blattenberger, & Verlerger, 1976; Wilson, 1971). These studies vary significantly with respect to the endogenous variables modeled, the exogenous variables incorporated, and the statistical techniques used. Understanding the variation in each of these factors is critical for understanding the differences among the conclusions reached. The variation in statistical techniques includes variation in the specification of relations, specification of error terms, aggregation of sampling units, optimality of estimators, and types of sampling frames (longitudinal versus cross-sectional). These issues, although of interest to statisticians, are not critical to the theme being developed since they are issues internal to the confirmatory approach. Addressed, instead, are the common elements of these studies, elements which are characteristic of the use of classical statistical methods to examine the demand for economic goods. These elements include the genre of techniques used, the philosophy which sustains the use of these techniques, and the strategy used in choosing which technique to apply. For the sake of analysis, I assume that techniques are not commonly used in the economic analysis of residential energy demand if they appeared in none of the studies reviewed.

As part of a related effort, a dozen studies (other than those listed above) which provide at least some analysis of energy demand in the residential sector were examined (Mayer, 1979a). It is important to note that some of the techniques of exploratory data analysis, such as the analysis of residuals from a regression model, were used in one or at most a few of these studies. On the other hand, these techniques are used, when they are used, as subsidiaries to classical statistical techniques and, as such, are part of the confirmatory mode. The philosophy and strategies central to the exploratory mode, however, did not appear in any of the studies.

Every analysis utilized one or more of the following classical methods: estimating the parameters of an *a priori* specified linear regression model using raw or transformed data and ordinary least squares or one of its econometric alternatives; normal theory hypothesis testing, particularly the testing associated with the normal theory regression model; estimating the parameters of probabilistic models by maximum-likelihood methods; estimating parameters of simple parametric time series modeling, usually autoregressive models; and fitting linear models to pooled longitudinal and cross-sectional data using variance-components techniques or similar techniques.

Being confirmatory studies, the strategy of inquiry followed in every study is as follows:

(i) Prior to looking at the data, some variant of the theory of consumer demand is embraced.

(ii) A set of dependent and predictor variables is chosen. The choice is based, in part, on the particular demand theory adopted; in part, on previous experiences with similar problems; and, in part, on the availability of data.

(iii) Again, prior to looking at the data, a mathematical model is developed. The form of the model is dictated, primarily, by consideration of the interpretability of the coefficients in terms of demand elasticities, and the sample frame used (longitudinal, cross-sectional, combined cross-sectional and longitudinal).

(iv) Assumptions are made about the stochastic behavior of the variables included in the model.

(v) The coefficients, or parameters, of the model are estimated by applying one or more of the standard econometric estimation techniques. The justification for using these techniques is that the estimators obtained are "optimal" in the sense that they display certain mathematical properties, such as unbiasedness, consistency, efficiency and asymptotic normality, provided the model and assumptions are correctly specified.

(vi) Hypotheses about the goodness of fit of the model and the sizes of one or more of the coefficients of the model are tested.

(vii) If the signs, the absolute sizes, or the relative sizes of the estimated coefficients are contrary to the underlying empirical theory, the model is usually reformulated. Reformulation may involve deleting variables from the model, adding variables to the model, changing the functional form of the model, or altering the assumptions about the stochastic behavior of the variables in the model.

(viii) Assuming the model obtained is satisfactory, substantive inferences are made. These inferences follow from interpretation of the signs, absolute sizes, and relative sizes of the estimated coefficients, interpretation of the hypothesis tests, and interpretation of the goodness of fit of the model.

In summary, for understanding the determinants of residential energy demand, the studies reviewed all adopted some variant of the following approach: an empirical theory is embraced and operationalized by the selection of variables, the formulation of the model, and the adoption of assumptions; the coefficients of the model are estimated from the data; hypotheses are tested; and the coefficients of the model and the goodness of fit of the model are interpreted.

The scientific basis of this classical statistical approach can be characterized as one in which theories are adopted, models are developed, empirical data is confronted, and the theory is rejected or revised if the model is not consistent with the data. The analysts explored the data by adopting a theory, developing a model, and then by assessing the evidence against the model. Theory and models are improved incrementally by confrontation with data,

For the sake of contrast, suppose that the analyst believes that because of imperfections in the markets for residential fuels, the physical factors affecting the demand for such fuels, and the sociological and political determinants of demand, the available theories are not *a priori* suitable for adoption. He or she might consider exploring the data without adopting a model and thus might proceed in the exploratory mode. No exploratory analysis of residential demand, other than those due, in part, to the author, could be found (Horowitz & Mayer, 1977; Mayer, 1978b, 1979b; Mayer & Benjamini, 1978; Mayer & Horowitz, 1979). Tukey (1977) gives a brief example using a subset of the data in Mayer (1978b).

2. The Strengths and Limitations of the Modes of Analysis

To be proficient, the data analyst should be intimately familiar with the benefits and the limitations of working in each mode. Deciding which mode to adopt for approaching a particular substantive problem is facilitated by familiarity with the characteristics of each mode. It should be emphasized that the modes are noncompetitive in the sense that they require different types and levels of theoretical underpinnings and yield different types and levels of scientific inference. Put simply, the confirmatory mode requires elaborate statistical assumptions which follow from a theory of the behavior of the process being studied and yields both statistical statements about the degree of accuracy of parameter estimates and probabilistic statements regarding the degree of evidence against certain hypotheses. The exploratory mode requires little or no statistical assumptions and no formal empirical theory and yields statements which are neither inferential nor probabilistic. It offers no formal mechanism for making inferences about the population from which the sample was generated. The rough confirmatory mode requires that a model be tentatively, albeit skeptically, adopted. It yields the inferences of the confirmatory mode as intermediate products and then relinquishes any statistical assumptions to explore the data. The final output is both confirmatory and exploratory in nature but is heavily influenced by the model and the associated assumptions adopted.

The major strength of the confirmatory mode is that the results obtained are replicable and optimal in well-defined areas. Given the same theory, models, and hypotheses, two investigators should obtain the same parameter estimates and the same degree of evidence in favor of each research hypothesis. Ideally, their results would differ only to the extent that they give different substantive interpretations to the statistical inferences made. Furthermore, should they offer different statistical results, then the optimality theorems of classical statistics often provide justification for preferring one set of results over the other. In areas such as multivariate analysis of variance, where several testing procedures compete but few optimality results obtain, the optimality and replicability of the inferences made are reduced and, consequently, the power of the confirmatory mode is impaired. In such areas statisticians are actively searching for appropriate optimality theorems.

The second strength of the confirmatory mode is that it yields indicators of the uncertainty associated with its inferences. As examples, the estimated standard error is an indicator of the uncertainty of a parameter estimate and the critical level and power are indicators of the uncertainty associated with a hypothesis test. Without the confirmatory mode the data analyst might be able to make inferences about the population being sampled, but these inferences would be informal in the sense that no indicator of the uncertainty of the inferences could be given. Such inferences would be significantly weaker, in the scientific sense, then the inferences of the confirmatory mode.

The major limitations of the confirmatory mode are that it requires specification of the theory, models, and hypotheses prior to the examination of the data. Confrontation with data allows the analyst to modify the model and propose alternative hypotheses, but classical statistical methods are designed to provide optimal estimators and tests and are not specifically designed to be aids in the reformulation of the theory, models, or hypotheses. Once the model and hypotheses have been changed, confirmatory methods are not optimal for estimating the model or testing the hypotheses. In fact, there are almost no theorems in classical statistics which guarantee the optimality of a confirmatory method if the model or hypotheses are reformulated after an initial confrontation with the data.

The major strength of the exploratory mode is that it permits the unstructured examination of data and transforms the act of data analysis from a scientific search for mathematical results into an artistic search for empirical patterns. It permits any manipulation of the data, but produces results that are simple to understand. Exploratory techniques may involve sophisticated computer routines, but never involve elaborate mathematical results. The exploratory mode brings the analyst close to the data since the results arise from the mind. Data structure is both exposed and imposed by the analyst. The inferences made by the analyst are informal and open to criticism. Each analyst must reach his or her own conclusions and these conclusions are open to question by any other analyst.

The major limitation of the exploratory mode is that it does not yield statements which are inferential in the statistical sense or replicable in the scientific sense. Although the analyst uncovers patterns in the data which are suggestive of hypotheses that could be tested in future studies, the analyst is not free to test these hypotheses on the data used to generate them. Consequently, the exploratory mode is best viewed as the first step in a sequential analytic process. It suggests hypotheses which are tested in subsequent analyses. These hypotheses are often improved by rough confirmatory analysis and finally tested by the confirmatory mode. Thus, unlike the confirmatory mode, the exploratory mode cannot stand alone as a method of inquiry.

A second limitation of the exploratory mode is that it usually benefits from more extensive data than that required by the confirmatory mode. Provided certain assumptions are made, the confirmatory mode can study aggregate units in order to model the behavior of individual sampling units (consumers). The exploratory mode has no provision for substituting aggregate for disaggregate data. Although the data on aggregates can be explored, the exploratory analyst prefers to examine data on individual units if the behavior of the individuals is the topic of analysis. This preference often leads to the gathering and management of a large amount of data.

The strength and weaknesses of the rough confirmatory mode can be gleaned from the strength and weaknesses of the modes it combines.

Thus far the limitations of the confirmatory and exploratory modes have been broached under the assumption that the canons of the mode are correctly followed. Additional limitations arise from the fact that a mode can be adopted in name but violated in spirit. For the confirmatory mode the most common violation arises from the fact that operating in the confirmatory mode encourages the analyst to modify the model incrementally and rarely leads to challenges of the underlying theory. It is easier to modify the model to fit the data than to reject the underlying theory. Theories tend not to be rejected and thus appear to attain an exalted position well above the data.

In a recent review, a senior economist suggested that

> It also proved to be possible to build models that were equally good from a number of different perspectives. Theories could not be accepted or rejected based on the data ... The data simply were not powerful enough to test and to choose among theories.
>
> As a result, econometrics shifted from being a tool for testing theories to being a tool for exhibiting theories. It became a descriptive language rather than a testing tool. Statistical models are built to show that particular theories are consistent with the data. But other theories are also consistent with the data and only occasionally can a theory be rejected because of the data. As a result good economic theory was stronger than the data ... and therefore it must be imposed on the data. What started out as being a technique for elevating data relative to theory ended up doing exactly the opposite (Thurow, 1977, p. 83).

The tendency to put theory before data is endemic to the confirmatory mode as applied in all areas. Although classical statistical methods can be used to formulate or reformulate a model from the patterns found in the data, they are not designed for that purpose. Neither is the philosophy or strategy of the confirmatory mode oriented toward model formulation. In fact, texts on classical statistics and econometrics rarely mention the problem of using data to formulate a model. Topics central to model reformulation, such as the systematic examination of residuals and robust estimators, are systematically ignored. Topics such as the optimality of estimators and the power of tests, topics which are covered in texts on classical statistics, hold the confirmatory spotlight. It is reasonable to suggest that methods designed to provide optimal statistical functions and objective probability statements may be less than effective in diagnosing the inadequacy of a theory or model.

For the exploratory mode, the most tempting violation is to make statistical inferences about the population being sampled. The analyst working in this mode must take care to explain that the models obtained are suggestive but have not been scrutinized by the rigor of confirmatory analysis. There is a constant temptation to test the model on the same data used to generate the model, a serious violation of the exploratory philosophy. The analyst is free to compute any hypothesis tests or estimators but is not free to report any statements which have statistical or probabilistic content. The decision to adopt the exploratory mode is a decision which proscribes the use of classical inferences before or after the exploratory analysis is complete.

A final comment is that an analysis in the exploratory mode can involve any methods of analysis, including those of classical statistics, provided no formal inferences are made about the population at hand. The exploratory analyst will probably want to use some techniques designed to explore data but is perfectly free to use techniques designed for the confirmatory mode. In fact, the analyst trained in the confirmatory mode and trying to work in the exploratory mode may be most comfortable, particularly in the early stages, using the methods of classical statistics—less the statistical inferences—to explore data.

3. The Confirmatory Analysis of Residential Energy Consumption

Confirmatory analysis of the demand for fuels in the residential sector is well exemplified by the widely cited study of Mount *et al.* (1973). The main objective of the study is "to measure the relationship between the demand for electricity and causal factors such as price (Mount *et al.*, 1973, p. 1)." The authors begin by presenting, without theoretical justification, a model of electricity demand. "Electricity demand [at the state level] is assumed to be determined by five factors: population, income, and the prices of electricity, substitute fuels such as [natural] gas, and complementary products such as household appliances (Mount *et al.*, 1973, p. 1)." Based on this assumption, the authors specify three models of electricity demand, the simplest of which is

$$Q_{it} = A_k Q_{it-1}^{\lambda} V_{1it}^{\beta_1} \cdots V_{Nit}^{\beta_N} \varepsilon \tag{1}$$

where Q_{it} is the quantity of electricity demanded for state i and year t in the residential, commercial, or industrial sector, λ, $\beta_1, \ldots, \beta_N$ are unknown parameters, ε is an unknown error term, A_k an unknown parameter corresponding to the kth region of the United States, and V_{jit} the value of the jth explanatory variable for the state i and year t.

The other two models specified are similar to the model displayed in (1). Each is assumed to apply to demand for electricity in each sector. The geometric lag structure of the model is defended in that the parameters are easily interpreted: β_j is the short-run elasticity of demand for the state relative to exploratory variable j, $\beta_j/(1 - \lambda)$ is the long-run elasticity, and $1 - \lambda$ is the proportion of the state's demand response in the first year.

The data used to estimate the parameters of the model are pooled cross-sectional and time series data for 47 states from 1946 to 1970. Although methods are available for estimating the parameters of pooled cross-sectional and time-series data, the authors assume that the lagged endogenous variables and the error variables are (asymptotically) uncorrelated and then utilize ordinary least squares (OLS) procedures. In case the assumption of uncorrelatedness is false, the authors also compute instrumental variable (IV) estimates, estimates which are consistent even if the assumption of uncorrelatedness does not hold.

By using ordinary least squares and instrumental variables, the authors assume that the structure of the data is known, up to random error, and that assumptions can be made about the behavior of the random error. The primary role of the data is in estimating the size of the coefficients of the model and testing the significance of these coefficients.

The coefficients obtained by applying OLS or IV are examined and then restrictions are imposed on the coefficients, when it appears desirable, "on statistical grounds and also to maintain the economic logic of the estimated elasticity values (Mount *et al.*, 1973, p. 8)." For example, the sample coefficient of log income in two of the models was negative and thus was constrained to be zero in the final model to ensure "that the income elasticities are nonnegative at high income levels (Mount *et al.*, 1973, p. 8)."

The final models obtained have squared correlation of over .99; no analysis of the residuals of the models is mentioned and no residual plots or diagnostics are displayed.

The estimated coefficients, t values, and long-run elasticities for the model displayed in (1) are given in Table 1.

The squared multiple correlation for the model is over .99.

As may be appropriate for a confirmatory analysis displaying a high degree of correlation, the models obtained are embraced as the correct models of electricity demand. Consequently, the conclusions presented by the authors stem from direct interpretation of the estimated coefficients and hypothesis tests. For example, "the major conclusion of the analysis is that the price of electricity is more important than income in terms of the long run elasticity (Mount *et al.*, 1973, p. 8)." The implications suggested for the regression coefficients and functions of the coefficients are far reaching and profound. For example, "price also plays a role in determining the life style of residential consumers (e.g., whether or not to install air conditioning) and the types of facilities and production methods employed by consumers and industrial consumers. However, it should be remembered that demand responds relatively slowly to changes in the causal factors . . . If prices increase over the next few years in response to increased fuel costs, etc., the growth of electricity demand will gradually decrease from the present rate (Mount *et al.*, 1973, p. 11)."

The study displays two sets of reservations about the interpretation of its statistical results. The first set pertains to abuses of the canons of the confirmatory mode. First, there is no economic theory of the behavior of states. The reader cannot question the derivation of the model when no theoretical support for the model is given. The theory on which the model is probably loosely based, the theory of consumer demand, does not apply directly to the behavior of states. If it is being adapted to apply to states, the adaptation must be stated if it is to be open to challenge. Second, if a theory

TABLE 1

	Estimated coefficient	t value	Long run elasticity
Lagged demand	.8859	136.5	—
Population	.1075	17.0	.94
Income	.0342	4.2	.30
Price of electricity	.1385	12.5	−1.21
Price of gas	.0238	6.7	.21
Price of appliances	−.0408	2.4	−.36
Constant by region (for three regions)			
North East	.4620		
Mid Atlantic	.4612		
East North Central	.4802		

of the behavior of states is being embraced tacitly, the study should, but does not, allow the data to challenge the theory. For example, when the data indicate results counter to the model, the model is modified and reestimated but no theory is ever doubted, a violation of the confirmatory mode.

Third, three competing models for the same response are estimated from a single data set with no single model being embraced, a good exploratory strategy, but one that does not belong in a confirmatory study. Fourth, two methods are used to estimate the parameters of the models and inferences are made from both. Such duplication violates the confirmatory approach since both sets of inferences may be seriously biased.

Fifth, the functional forms of the models presented are never justified or related to theory. They are motivated by ease of interpretation, a weak justification for confirmatory models. Sixth, if the parameter estimates differ from the authors' expectations, the expectations are preserved by setting the parameters equal to zero and reestimating the model. Inferences about models are made without regard to the rejection of the first model. Seventh, the confirmatory structure of the study involves modeling the behavior of states, but some of the interpretations, given the results, are about the behavior of individual consumers. The inferences do not appear to follow from the model developed. The authors do not present enough results to defend or challenge their linkage between the behaviors of states and individuals. For example, the conclusion that consumers reduce their demand in response to increase in price cannot follow immediately from a model of state level demand. Finally, some of the conclusions made, such as conclusions about the life styles of residential consumers, appear not to follow whatsoever from the statistical results presented. Confirmatory analysts need to distinguish between that which is being confirmed and that which is speculation on the part of the analyst.

One could eliminate many of the weaknesses presented by assessing the study as a rough confirmatory analysis. Unfortunately, the study has serious weaknesses if assessed as such. First, and most critically, it fails to analyze deviations about the models developed. Although the correlations obtained are high, systematic deviations from the model could exist which, when explored, might indicate improvements in the model. Second, additional predictor variables are not considered after the model is fitted. Third, a wide range of functional forms for the models are not entertained. Fourth, residuals from competing models are not compared in order to indicate which model is superior.

Fifth, possible changes in characteristics of the conditional distribution of demand other than the conditional mean are ignored. For example, the exogenous factors might affect the variance of the demand. Sixth, the joint distributions of two or more variables are never examined. Seventh, the simultaneity assumption required to use a single equation model is never tested. Eighth, alternative lag structures are not tried. Finally, the bias of using cross-sectional data to make dynamic inferences is not assessed.

The study fails to satisfy the guidelines of either a confirmatory or a rough confirmatory analysis. The problem may be that too little is known about the demand of interest to allow the analyst to proceed in either mode. In the next section an exploratory analysis of the demand for residential fuels is given.

A final comment is that none of the studies reviewed assessed data problems such as missing data or the possible lack of reliability of the data. Confirmatory studies tend to require complete, reliable data sets and yet energy data tend to be fraught with inadequacies. The data problems faced in any statistical study are worthy of mention.

4. An Exploratory Analysis of Residential Energy Demand

An exploratory analysis of the relationship between the price of and demand for natural gas in the residential sector is summarized by Horowitz & Mayer (1977). This study is one component of a larger statistical effort which looked at various aspects of the demand for energy in a single planned residential community (e.g., Mayer 1978b, 1979b; Mayer & Benjamini, 1978; Mayer & Horowitz, 1979). The studies all use disaggregate time series data in which the unit of analysis is the individual household. The object of the analysis is to explore data on natural gas consumption to uncover patterns which are suggestive of models and hypotheses worthy of further investigation. No inferences are made about the factors which influence demand for fuels in the general population of consumers.

The data on which the analysis reviewed focuses are the monthly natural gas consumption statistics for 1971 to 1976 for 401 owner-occupied townhouses. The data were collected as part of a complex interdisciplinary study of the physical and behavioral parameters of the energy environment of a single residential community (many of the results of the study are presented in Socolow, 1978b, and summarized in Socolow, 1978a). The community, Twin Rivers, New Jersey, is located about 45 miles from New York City. The units being studied are of only two styles. There are 248 two-floor apartments and 153 split-level apartments which were constructed by a single builder and are of a single vintage, having been built in 1970. Natural gas is used solely for space heating; the water heaters, air conditioners, and appliances are electric.

As is typical of an exploratory analysis, the study has a somewhat narrow focus. It uses disaggregated data to examine the demand for a single fuel in a sample of almost identical homes, served by a single utility, facing a single price schedule, where the fuel of interest is used for a single purpose. The results reviewed here are for the two-floor units, but the only patterns which are reported are those which were also found in the data on split-level units, the two subsets being used for informal cross-validation of all results.

The only major variations in the design among the two-floor units are the number of bedrooms (and thus the size of the unit) and the number of neighbors (one or two). The community is unusually homogeneous demographically. The families tend to be white middle-class natives of the New York region. Heads of households tend to be about 30, college educated, with one or two children, and family incomes in the $25,000–30,000 range. The lack of variation in income permits the examination of the relationship between price and demand without consideration of income.

The units which had a change in occupants, stood empty for a long period, had incomplete records, or had obviously "bad" data were removed from the sample since the goal is to estimate the response of households to changes in price; the sample was thus reduced to 151 units.

The analysis begins by examining the price and cost of natural gas. The Twin Rivers consumers face a typical per-therm block pricing system with a highly discontinuous base price and a per-therm fuel adjustment charge which is highly variable in time. After considering several alternatives, the marginal price was operationalized as the price of a therm for consumption over 50 therms, 50 therms being the final step on the block pricing system. The marginal price for the winter months is displayed in Figs. 1 and 2 in current and constant dollars. Over the period of the study, the marginal

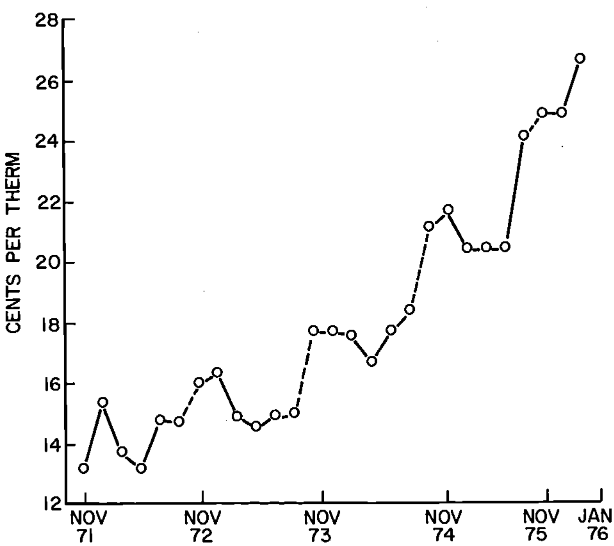

Fig. 1. Marginal price of natural gas at Twin Rivers: current dollars.

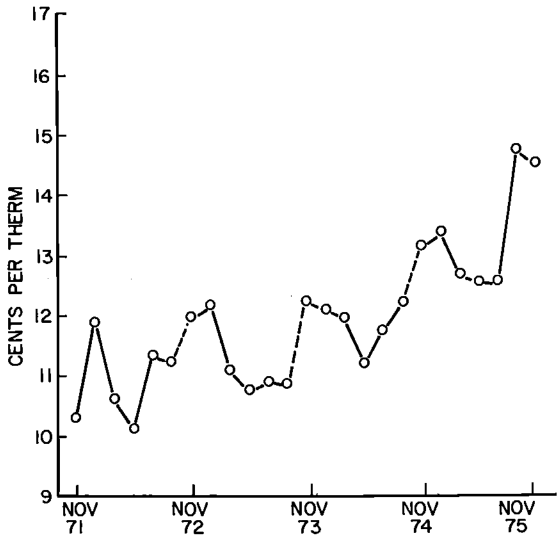

Fig. 2. Marginal price of natural gas at Twin Rivers: 1967 dollars.

price rose from 12 to 28 cents in current dollars, and from 11 to 15 cents in constant dollars. The first pattern which may be worthy of note is that, due to the structure of the fuel adjustment charge, the marginal price peaks in early winter (November and December) and then decreases for the coldest months (January and February).

The marginal price for the nonwinter months is not shown although the demand for these months shows several notable patterns. The demand in the summer, for example, averages only 8 therms per month, that being due to the pilot light of the furnace. The consumers, however, pay a flat minimum charge for small demand. Thus consumers have no economic incentive to reduce their summer demand, possibly by closing their pilot lights.

Most of the econometric studies use average prices for an aggregate of consumers and not marginal prices to predict the demand for fuels. The choice of average price is usually one of convenience and not theory. Marginal price is difficult to operationalize when modeling aggregate consumers. Figure 3 displays the actual average cost per therm (multiplied by 150 for convenience—150 therms being the average demand for natural gas in a winter month). This variable appears to be more sensitive to changes in the rate structure than is marginal price. Note that changes in the structure can significantly affect the cost per therm without affecting the marginal price. Due to the rate structure, the colder the winter the less the average cost per therm. The average monthly cost (bill) divided by the number of degree days (a simple measure of the coldness of the month) is displayed in Fig. 4.

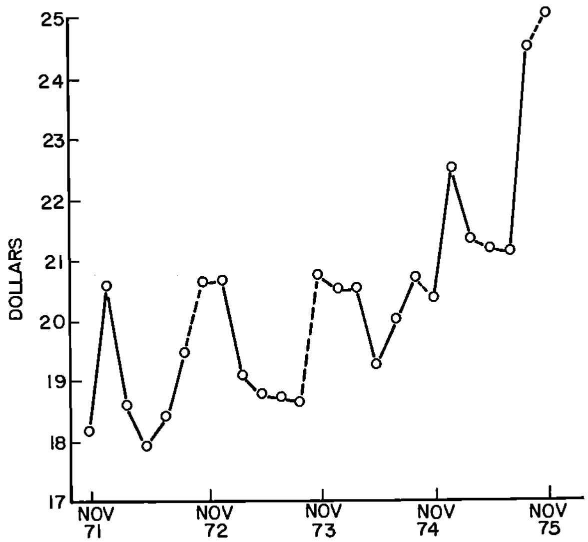

Fig. 3. Average cost per therm (× 150) of natural gas at Twin Rivers.

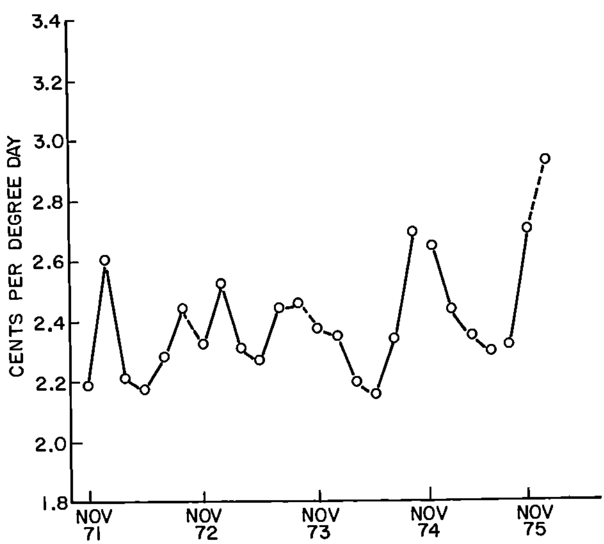

Fig. 4. Average cost per degree day at Twin Rivers.

Comparisons of Figs. 2–4 indicate that the marginal price, cost per therm, and the cost per degree day are not as similar as might be expected. It is interesting to speculate whether consumers respond more to the marginal price, the cost per therm, or the cost per degree day for purchases of residential heating fuel. We invite economists, and others, to speculate.

The differences among the price variables are due, in part, to frequent changes in the structure as well as the rates of the price schedule. A search of the literature on energy economics revealed no theory for how consumers respond to such changes.

A second view of the cost of natural gas is presented in Figs. 5 and 6 which display the comparison boxplots (Tukey, 1977; McNeil, 1977) of the monthly cost (bills) and the monthly cost per degree day, respectively, for natural gas for the individual townhouses. These plots show a large amount of variation in demand among consumers living in almost identical units, a variation that is ignored when aggregate units are modeled. Among the patterns worthy

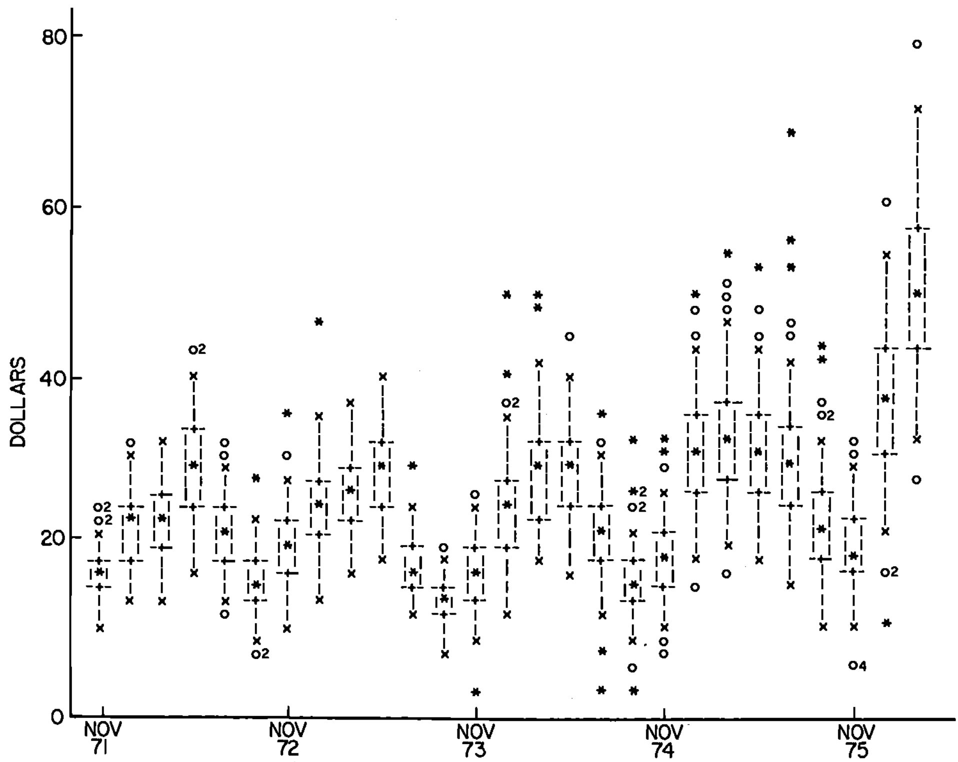

Fig. 5. Distribution of the monthly cost of natural gas.

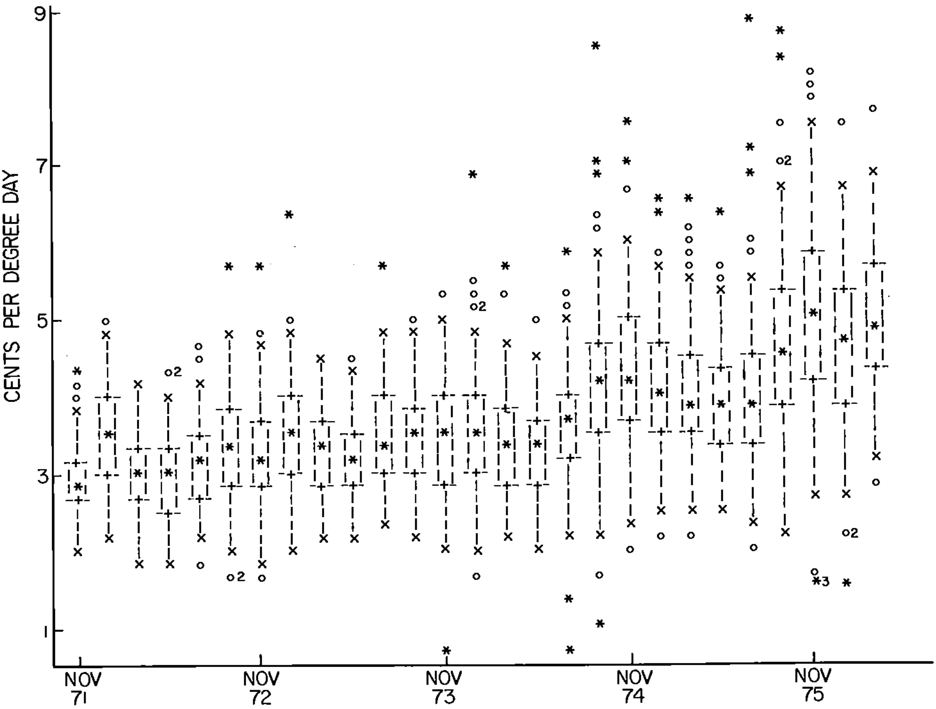

Fig. 6. Distribution of the cost per degree day at Twin Rivers.

of note are the following:

(i) The average January bill rose, in five years, from \$22 to \$50 and the largest January bill rose from \$33 to \$79.

(ii) Large increases in the monthly bills occurred for the coldest months but only small increases appeared for the mildest winter months.

(iii) The range of monthly cost increased dramatically, going from \$25 for January 1972 to over \$50 for January 1977.

(iv) The range of cost in any given month is surprisingly large considering that the units are almost identical and gas is used solely for space-heating.

(v) The distributions of monthly bills tend to be symmetric with slightly more extreme values than would be expected from Gaussian distributions.

(vi) The cost per degree day increased, on the median, by about 60%.

The variation in demand among almost identical units may be due to differences in income, appliance stock, demographic characteristics, the quality of construction, and the behavioral patterns of the occupants (Mayer, 1978b).

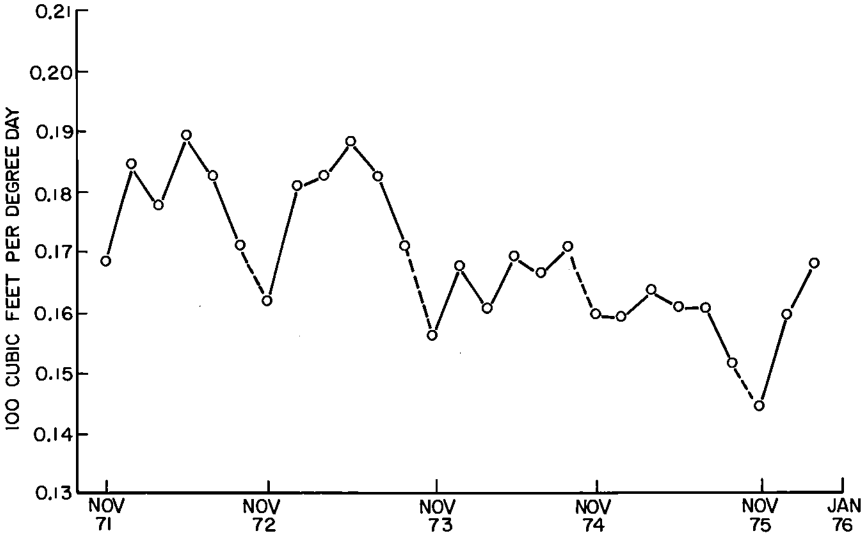

Fig. 7. Average monthly demand for natural gas at Twin Rivers.

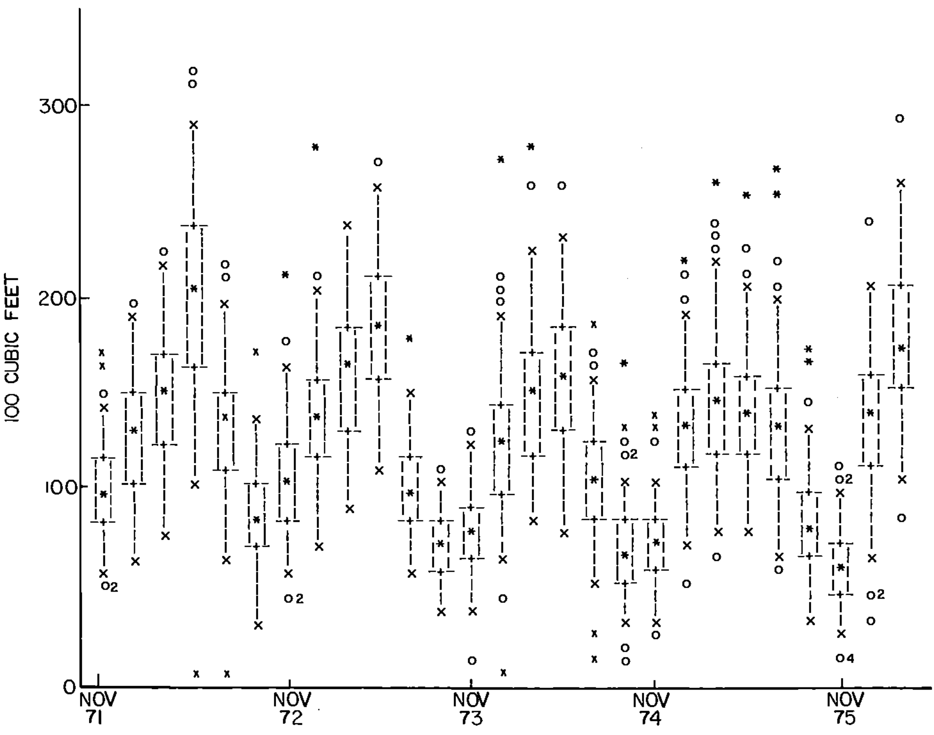

Fig. 8. Distribution of the monthly demand for natural gas at Twin Rivers.

Understanding the variation in cost could be an important aid in developing energy policy since policies which reduce the highest consumers but maintain the lowest may be as wise and as practical as policies which only affect the mean.

The second category of variable of interest, demand, is considered with average demand per degree day displayed in Fig. 7 and comparison boxplots of monthly demand presented in Fig. 8. As expected, patterns similar to those for the cost variables are found: demand is highly dependent on the calendar month, it appears to have decreased after the onset of the Arab oil embargo (cf. Mayer, 1978b), the range has increased, and the distributions of monthly demand are symmetric with numerous extreme values. Degree days for each month are displayed in Table 2.

The variation in demand is surprisingly large. Our efforts have shown that the ratio of the highest demand to the lowest demand in a single month for our sample is over 3 to 1 after vacancies and vacations have been eliminated. For units with the same number of bedrooms, neighbors, and similar appliance stocks, the ratio is over 2 to 1. We suggest this variation is worthy of further study and speculation. It is important to note that it is outside the dominion of any model of aggregate demand.

Since the demand for natural gas is heavily dependent on the coldness of the month, it is necessary to adjust the demand figures to account for differences in weather. Several indicators of demand adjusted for weather were explored, including demand per degree day and demand adjusted for degree days, and the residual variable obtained for each unit from the regression of monthly demand on degree days for that unit.

Both indicators are plagued by problems. The first ignores the fact that some portion of the demand is for the pilot light and that portion does not depend proportionally on degree days.

The second indicator has two problems. First, one or two months are extreme values with regard to degree days and these extreme values heavily

TABLE 2

THE NUMBER OF DEGREE DAYS FOR EACH MONTH

	1971–1972	1972–1973	1973–1974	1974–1975	1975–1976
November	602	647	510	436	391
December	673	761	746	797	852
January	824	882	932	863	1077
February	1068	975	942	865	—
March	726	533	631	801	—
April	489	397	391	509	—

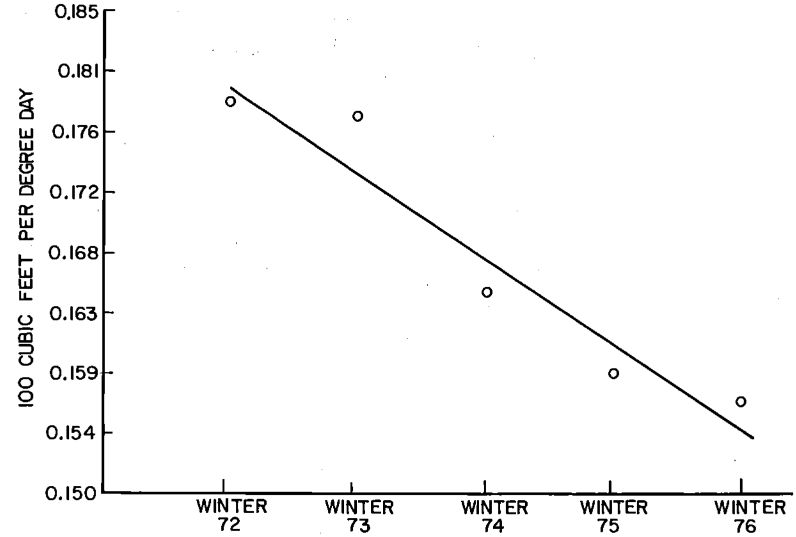

Fig. 9. Average demand per degree day for each year.

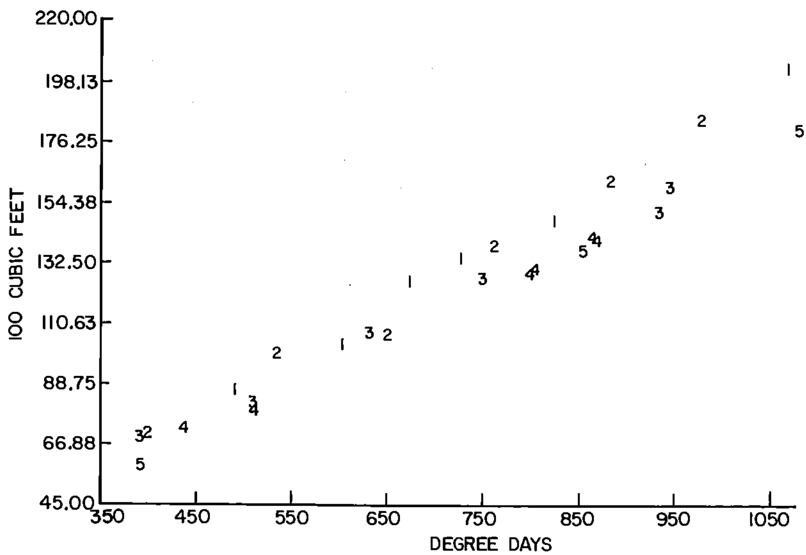

Fig. 10. Average monthly demand versus degree days.

influence the regression line. This problem was reduced by substituting robust regressions for least squares regressions (Mosteller & Tukey, 1977). The former are less influenced by extreme scores. Second, the relationship between demand and degree days was significantly altered by the onset of the Arab oil embargo. To show the alteration, average demand per degree day for each year is displayed in Fig. 9 and the average monthly demand is plotted against degree days with the year indicated in Fig. 10. Both figures indicate that a single regression cannot summarize the aggregate relation between demand and degree days. Figure 11 displays a time plot of the

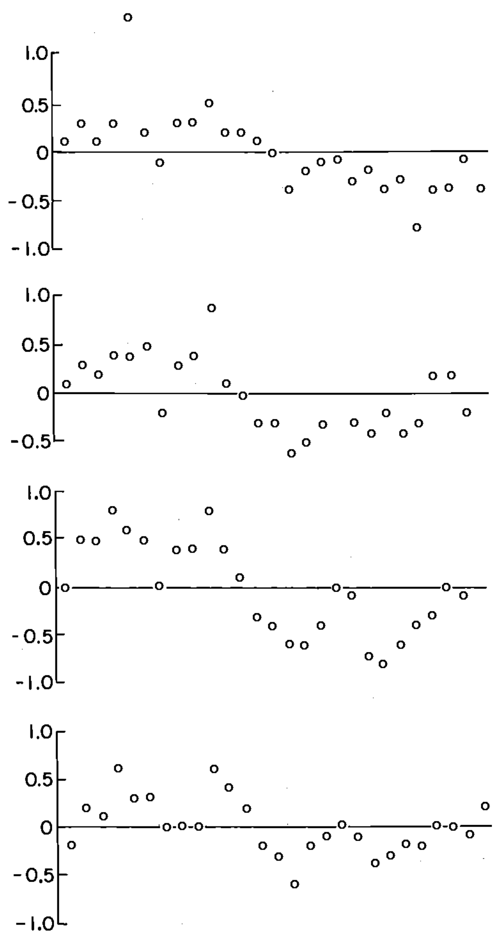

Fig. 11. The residuals of demand regressed on degree days versus time for several townhouses.

residuals of demand regressed on degree days for several individual units. Clearly the residuals are not random. In fact, analysis showed that in spite of the high correlation between average demand and degree days, no simple linear method of adjusting demand for degree days was adequate when all years were considered. We chose to adjust demand for degree days by robust regression using only data collected after the oil embargo.

The exploratory analysis of the relationship between price and demand used various graphical techniques to examine the characteristics of the joint distributions of the price variables and the demand variables. The major result can be summarized by comparing Figs. 2 and 7. Although demand decreases and price increases over the period of analysis, the major decrease in demand, a decrease of 13%, occurred in the winter of 1974 while the major increase in price occurred in the winters of 1975 and 1976. Thus there is little reason to believe that the most significant decrease in demand was driven by price. It may have been due to other economic factors, or it may be the response of Twin Rivers residents to the President's appeal to reduce their thermostats to 68°. Although several economists have argued with us that this decrease was due to price anticipation, none has been able to provide a convincing formal argument. Furthermore, using the available data there is no way to isolate systematically the cause of the dramatic 1974 decrease in demand.

Following the lead of the econometric studies but remaining in the exploratory mode, several models, including the following, were fitted to the data:

$$\log D = \alpha + \beta_1 \log P + \beta_2 \log DD + \beta_3 \log I + \varepsilon,$$

where D is average monthly demand, P is real marginal price, DD is the number of degree days, and I is an index of inflation. The estimated model using the data from 1971 to 1976 is

$$\log D = 19.6 + \underset{(.52)}{.65} \log P + \underset{(.0002)}{.0018} \log \mathrm{DD} - \underset{(.82)}{2.36} \log \mathrm{I}$$

with the standard errors of the coefficients in parentheses. Being in the exploratory mode no hypothesis testing is undertaken.

This model and similar models indicate that the effect of price on demand for the sample at hand is negligible and the effect of inflation is somewhat larger. Unfortunately, the data used covers the large 1974 decrease in demand. Residual plots show that the model, as expected, does not adequately fit this large demand. It became clear that no regression model can fit the data if it includes the dramatic 1974 decrease in demand.

Recently, in a separate study, models similar to the preceding were estimated, using data through 1978 but omitting data from the pre-embargo years (Tittmann, 1978). Although demand decreased each year during this period, the coefficients for both price and inflation were negligible. These efforts lead us to believe that for our sample price may have little effect on demand.

Turning to the variable demand per degree day, several models, including the following, were fit to the post-embargo data:

$$\log(D/DD) = \alpha + \beta \log P^*$$

where P^* is the price in current dollars. The resultant model is

$$\log(D/DD) = \underset{(.08)}{-2.063} - \underset{(.06)}{.22} \log P^*$$

which yields a correlation of $-.90$, the highest correlation between any indicators of price and demand. The coefficient $-.22$ might loosely be considered to be an elasticity of demand in current price. It indicates a fairly inelastic relationship.

All of the exploratory modeling was repeated on the sample of split-level units and similar results were found.

Subsequent analyses focused on assessing the short-term relationship between price and demand following the Arab oil embargo. In Fig. 12 demand adjusted for degree days is plotted against marginal price with the months indicated. Demand decreases as price increases with the major decrease in

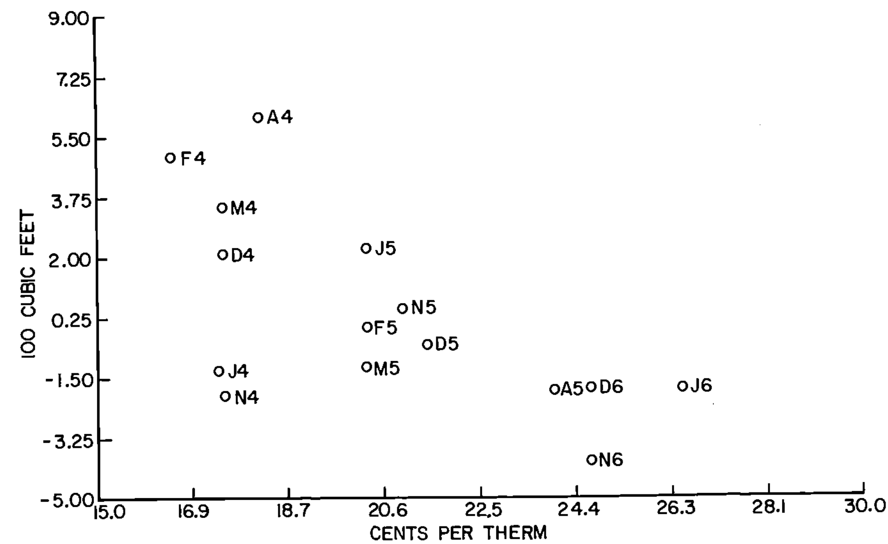

Fig. 12. Demand adjusted for degree days versus marginal price.

demand corresponding to years. To investigate further, log demand adjusted for degree days was regressed separately on log marginal price and years. The standard error of estimate for the regression of log demand adjusted for degree days on log marginal price is 2.20 with a correlation of −.67 and for the regression on years is 2.23 with a correlation of −.66. The relationship between demand adjusted for degree days and time is displayed in Fig. 13 and demand and years in Fig. 14. Demand is as structured in years as it is in marginal price. The same pattern holds when the data is updated through the winter of 1978. The analyses suggest that there has been some decrease in demand as a function of years but that demand does not fluctuate with marginal price within years.

In examining the joint distribution of marginal price and demand from 1974 to 1978 we noted an additional pattern: the real marginal price of natural gas at Twin Rivers *fell* slightly from January 1976 to December 1977, a 23-month period, but demand adjusted for degree days decreased steadily over this time span. Patterns such as this one gave us additional evidence that price may not drive demand in the short run for the sample at hand.

Operating in the exploratory mode, the data analyst usually tries multiple approaches to a single data set. Handicapped by the sacrifice of probabilistic inferences of the confirmatory mode, the analyst may want to "confirm" any results by using a second method of analysis. Usually the second analysis uncovers additional patterns as well as validates the original patterns. In

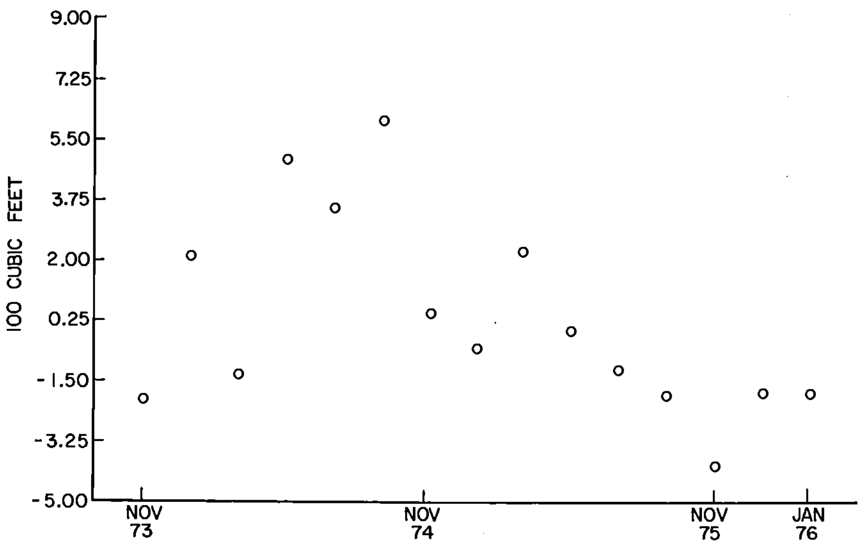

Fig. 13. Demand adjusted for degree days versus time.

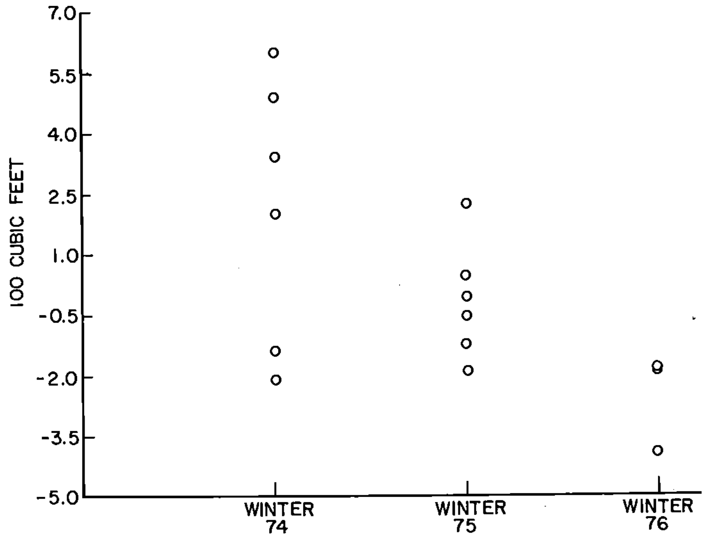

Fig. 14. Demand adjusted for degree days versus years.

addition to repeating the above analyses on the sample of split-level units, a second analysis was performed on both samples.

The variables of interest, marginal price, monthly cost, cost per degree day, price per therm, demand per degree day, and demand adjusted for degree days were decomposed by a robust additive analysis of variance model (McNeil, 1977; Tukey, 1977). The model yields total effects, year effects, month effects, and residuals for each variable. As an example, the year effects and month effects for demand per degree day are displayed in Figs. 15 and 16.

The year effects of demand and price, as expected, are only slightly related. The (calendar) month effects, as expected, show almost no relationship, indicating that controlling for years the fluctuations in demand and price over the calendar months are almost unrelated.

The residuals from the analysis of variance are the month to month fluctuations in the variables controlling for average demand, the year and the calendar month.

The residuals from the analysis of variance models tend to be symmetric in distribution with a few outliers caused by extremely warm or extremely cold months. The residuals for the demand variables are plotted against the residuals for the price variables and no patterns emerge. For example, the plot of the residuals of demand per degree day against the residuals of marginal price shows no pattern and produces a squared correlation of less than .1.

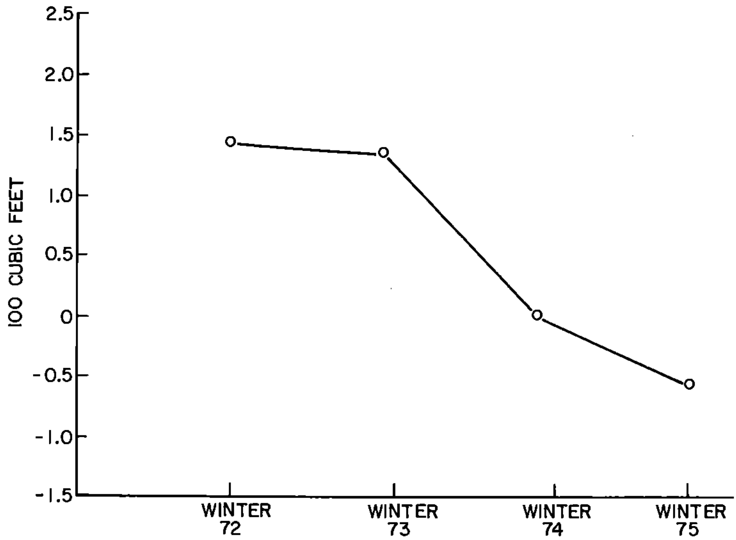

Fig. 15. Robust analysis of variance of demand per degree days: year effects.

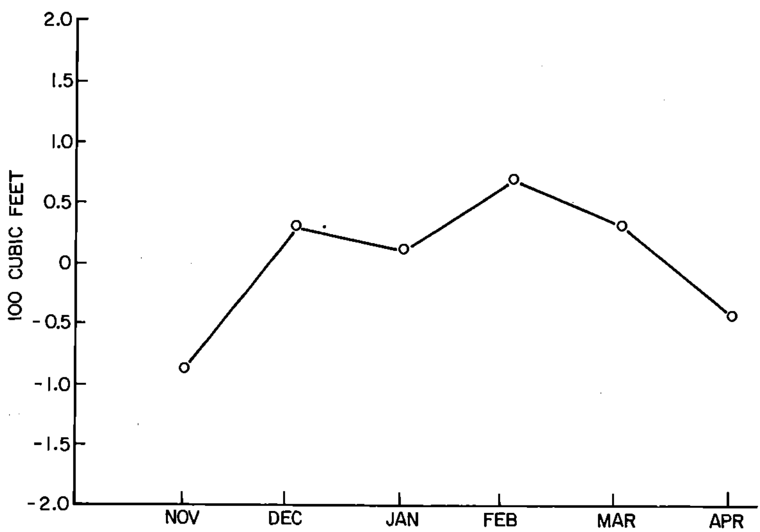

Fig. 16. Robust analysis of variance of demand per degree days: month effects.

This analysis and others suggest that controlling for the year and the calendar month, and adjusting for degree days, the fluctuations in demand from month to month are unrelated to the fluctuations in price.

Simple plots and regressions reveal that demand is explained by year as well as by marginal price and does not fluctuate with price within years.

Decomposing demand and price into total, year, and month effects and residuals and relating the effects and the residuals revealed no additional structure.

The exploratory analysis suggests that the relationship between the demand and price of natural gas for the sample being analyzed is complicated. While price has increased and demand has decreased over the last seven years, examination of these changes suggests that the major decrease in demand was not in direct response to the major increases in price. Over the entire period demand is more highly related to inflation than to the marginal price but is not highly related to either. Furthermore, no regression model could be found which fit the demand data well. The monthly fluctuations in price and demand appear unrelated.

The exploratory analysis, as expected, uncovers several patterns in the data worthy of further analysis. Among them are

(i) the large variation in demand among structurally and demographically similar townhouses;
(ii) the significant decrease in average demand which occurred after the onset of the Arab oil embargo;
(iii) the smaller decreases in demand which occurred in subsequent years;
(iv) the complex rate structure and the complex behavior of the marginal price of natural gas;
(v) the differences in the behavior of the various indicators of demand;
(vi) the effects of calendar months on the demand per degree day; and
(vii) the structure of the residuals of the regression of demand on price.

Additional analyses, possibly both exploratory and rough confirmatory in nature, could address these patterns and try to develop models of the relationship between price and demand.

5. Conclusions

The confirmatory and exploratory analyses of the residential demand for fuels are quite different. The former assumes a model, estimates parameters, tests hypotheses, and concludes that price was an important determinant of the residential demand for electricity when compared to other explanatory factors in the model.

The exploratory analysis explores disaggregate data at length and finds no model which adequately fits the demand for residential natural gas. It suggests several issues worthy of speculation and theory and concludes that for the sample at hand demand has decreased each year but there is not sufficient evidence that price drives demand in either the short or long run.

ACKNOWLEDGMENTS

The author gratefully acknowledges the support of the Office of Energy Programs, Department of Commerce, through Contract No. 6-35599 and the Office of Energy Information Validation, Energy Information Adminstration, through Contract No. EI-78-S-01-6540, awarded to the Department of Statistics, Princeton University.

The author also gratefully acknowledges the reviews of an earlier draft given by Jan Kmenta, Jim Ramsey, Kerry Smith, Katherine Bell, Robert Stine, and Scott Zeger.

REFERENCES

Anderson, K. P. The demand for electricity: Econometric estimates for California and the United States. Santa Monica, Cal.: Rand Corporation, 1972.

Baughman, M. L., & Joskow, P. L. Energy consumption and fuel choice by residential and commercial consumers in the United States. Unpublished rep., M.I.T., 1975.

Cohn, S., Hirst, E., & Jackson, J. Econometric analyses of household fuel demands. Oak Ridge, Tenn.: Oak Ridge National Laboratory, 1977.

Fisher, F. M., & Kaysen, C. *A study in econometrics: The demand for electricity in the United States.* Amsterdam: North-Holland Publ., 1962.

Griffin, J. M. A long-term forecasting model of U.S. electricity. *Bell Journal of Economics and Management*, 1974, **5** (2), 515–539.

Halvorsen, R. Residential demand for electric energy. *The Review of Economics and Statistics*, 1975, **57**, 12–18.

Hirst, E., Lin, W., & Cope, J. An engineering–economic model of residential energy use. Oak Ridge, Tenn.: Oak Ridge National Laboratory, 1976.

Hoaglin, D. C., & Welsch, R. E. The hat matrix in regression and ANOVA. *The American Statistician*, 1978, **32**, 17–22.

Horowitz, C. E., & Mayer, L. S. The relationship between the price and demand for natural gas: A partially controlled study. *Energy Research*, 1977, **2**, 193–222.

Mayer, L. S. Exploratory data analysis and classical statistics: Their ability to shed light on energy issues. In D. A. Gardiner & T. Truett (Eds.), *1977 Department of Energy Statistical Symposium Proceedings.* Vol. I. Oak Ridge, Tenn.: Oak Ridge National Laboratory, 1978. Pp. 27–32. (a)

Mayer, L. S. Estimating the effects of the onset of the energy crisis on residential energy demand. *Resources and Energy*, 1978, **1**, 57–92. (b)

Mayer, L. S. Econometric energy models: A critical essay. Technical Report, Department of Statistics, Princeton University, 1979. (a)

Mayer, L. S. The use of semi-controlled experiments in the analysis of residential energy demand. *Proceedings of the Department of Energy Statistical Symposium*, Washington, D.C.; Department of Energy, 1979, in press. (b)

Mayer, L. S., & Benjamini, Y. Modeling residential demand for natural gas as a function of the coldness of the month. In R. H. Socolow (Ed.) *Saving energy in the home.* Cambridge, Mass.: Ballinger, 1978. Pp. 301–312.

Mayer, L. S., & Horowitz, C. E. The effect of price on the residential demand for electricity: A statistical study. *Energy*, 1979, **4**, 87–99.

McNeil, D. *Interactive data analysis.* New York: Wiley, 1977.

Mosteller, F., & Tukey, J. W. *Data analysis and regression.* Reading, Mass.: Addison-Wesley, 1977.

Mount, T. D., Chapman, L. D., & Tyrrell, T. J. Electricity demand in the United States: An econometric analysis. Oak Ridge, Tenn.: Oak Ridge National Laboratory, 1973.

Popper, K. R. *The logic of scientific discovery.* New York: Harper and Row, 1968.

Socolow, R. H. The Twin Rivers program on energy conservation in housing: Highlights and conclusions. In R. H. Socolow (Ed.), *Saving energy in the home.* Cambridge, Mass.: Ballinger, 1978. Pp. 1–62. (a)

Socolow, R. H. (Ed.) *Saving energy in the home.* Cambridge, Mass.: Ballinger, 1978. (b)

Taylor, L. D., Blattenberger, G. R., & Verlerger, P. K., Jr. The residential demand for energy. Washington, D.C.: Electric Power Research Institute, 1976.

Thurow, L. C. Economics 1977. *Daedalus*, 1977, **106** (Fall), 79–94.

Tittmann, P. C. Conservation trends in natural gas consumption from winter 1974 to winter 1978. Junior Thesis, Department of Statistics, Princeton University, 1978.

Tukey, J. W. *Exploratory data analysis.* Reading, Mass.: Addison-Wesley, 1977.

Welsch, R. E. Regression sensitivity analysis. Unpublished manuscript, M.I.T., 1977.

Welsch, R. E., & Kuh, E. Linear regression diagnostics. New York: National Bureau of Economic Research, 1977.

Wilson, John. Residential demand for electricity. *Quarterly Review of Economics and Business*, 1971, **11**, 1–22.