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: Some Comments on the Evaluation of Informal Models

Chapter Author: V. Kerry Smith

Chapter URL: http://www.nber.org/chapters/c11696

Chapter pages in book: (p. 123 - 131)

Some Comments on the Evaluation of Informal Models*

V. KERRY SMITH

DEPARTMENT OF ECONOMICS
UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL
CHAPEL HILL, NORTH CAROLINA

The need to quantify the magnitude of economic impacts for all types of policy analyses together with the increased accessibility of computer software for a menu of econometric methods has served to hasten the diffusion of econometric analysis into diverse applications. Whereas in the past these applications were confined to areas where the analyst felt there was some established theory to guide the formulation of hypotheses, today the reach of econometric methods is both inside and outside the traditional realm of economic theory. Economists are involved in quantitative analyses of problems ranging from the analysis of the effects of air pollution on human health to simultaneous equation modeling of the effects of pronatalist and abortionist policies.[1] As a consequence, increasing attention must be given to the performance of existing methods and the development of new ones to meet the needs of these informal models. The diversity in the current and growing set of applications of econometrics therefore represents both a reflection of the power of existing econometric tools and an opportunity to understand them better in new settings.

The papers in Part I of this volume consider the different methodologies for dealing with informal models. For the most part they are techniques that are not widely known by applied econometricians and have been suggested

* These comments were prepared while the author was a staff member of Resources for the Future, Washington, D.C.

[1] For recent examples see Lave & Seskin (1977) and Coelen & McIntyre (1978).

ISBN 0-12-416550-8

here as potential additions to the methods used in empirical work. To facilitate this discussion, the next three sections will consider each theme individually. The last section briefly summarizes the remarks.

1. Exploratory Data Analysis

In his paper Mayer compares the relative usefulness of Tukey's (1977) suggestions for more intensive data analysis in many statistical applications. He argues persuasively for this approach rather than the adoption of specific, and often arbitrary, models in order to appeal to available theorems for hypothesis testing and parameter estimation. Mayer observes that the econometric approach to data analysis is similar to classical statistics in that it places too much emphasis on parameter estimation and hypothesis testing and therefore impedes the analyst from uncovering the subtleties of his (or her) data. Therefore he proposes the adoption of an organized approach to understanding the features of a data set, which is largely independent of any formal model. It is called exploratory data analysis and can operate in any of three modes: confirmatory, rough confirmatory, and exploratory. Following his clear discussion of each of these modes of analysis, Mayer uses a study by Mount, Chapman, & Tyrrell (1973) to illustrate the shortcomings of conventional econometric approaches to energy demand analysis. In addition, the paper uses more recent data, including a sample of individual household's consumption patterns of natural gas, to consider explicitly how an exploratory data analysis might proceed in assisting in the development of models for residential energy demands.

It is difficult to quarrel with the suggestion that analysts should attempt to understand fully the strengths and weaknesses of their data as an *integral* part of the process of using them to learn about the problem under study. However, Mayer's arguments seem to extend beyond this unassailable position to suggest that many (if not most) conventional econometric analyses of the demands for energy resources have failed to conform to this practice. His example (the Mount *et al.* analysis) is cited as one which abuses the "canons" of the confirmatory mode of analysis. Moreover, Mayer's overall summary might leave the unwary reader with the impression that the significant problems present in many energy data series have not been given serious consideration in these early econometric studies of demand. This characterization is misleading. Indeed, the impetus for the development of the new data sets, such as the one Mayer uses later in his paper, arose from the identification of data problems in these earlier studies.[2]

[2] See Halvorsen (1975), Taylor (1975), and Fisher & Kaysen (1962) as examples.

Nonetheless it is important to note that Mayer's criticisms of the Mount *et al.* paper raise more general questions with respect to the presentation of most applied econometric work. It is often difficult to judge the amount of data analysis, in a rough confirmatory mode, that has been undertaken with a given application on the basis of the published findings. While there are most certainly variations in the care exercised and depth of the analyses, the absence of reports on the details of the process undertaken does not necessarily imply an absence of the analysis itself. Restrictions on the length of publications and the like have often contributed to the omission of this information.

Similarly, misinterpretations can creep into abbreviated discussion of economic models. Consider, for example, Mayer's appraisal of the Mount *et al.* economic model. He does offer a number of useful insights, although there are also some serious difficulties with Mayer's understanding of these authors' objectives. Mayer has assumed that, because the Mount *et al.* sample is a pooled time series of states, the authors are attempting to model the demand for electricity *by a state.* This conclusion is incorrect. The analysis relates to the measurement of the aggregate demand for electricity by the customers within a given service class (i.e., residential, commercial, or industrial). While the study differs from most of the other studies using similar data bases in that other authors used quantity per customer as a measure of the average demand, both approaches encounter aggregation problems. Given the paucity of data at the time the study was undertaken the strategy of Mount *et al.* must be regarded as one possible compromise.

Unfortunately by the close of the third section of his paper Mayer left this reader with the impression that economists have shown little interest in analyzing the methods actually used to estimate economic models and that they have a more limited comprehension of when the problem at hand does not warrant a confirmatory mode. This conclusion would again be misleading. While the names are different, it does seem fair to suggest that there has been intensive interest in this area in the recent literature under various "catch words" such as specification searches, data-mining, pretesting, sequential estimation, and the like. One can find evidence of concern nearly 20 years ago.[3] For example, in reviewing the problems associated with the application of econometrics, Theil (1961) proposed that estimates must be considered to be the result of the full strategy used in arriving at them. Moreover, to the extent that this strategy involved using a single sample to refine a model step by step, we can expect that if the model's estimated precision is based on the conventional prescriptions of classical inference then

[3] There is extensive literature in this area. Two useful reviews have been conducted by Judge, Bock, & Yancey (1974) and Wallace (1977). Leamer (1978) has also developed the analysis from the Bayesian perspective.

it will be biased. More specifically, Theil noted that

> Economic theory can give some indications as to the variables that are possibly relevant; it may even give some vague indications as to curvature and as to the numerical magnitude of some coefficients . . . but it rarely gives any indication about probability properties of disturbances. The obvious result is that, if a "maintained" hypothesis gives unsatisfactory results, it is not maintained but rejected, and replaced by another maintained hypothesis; etc. It is hardly reasonable to say that this kind of experimentation is incorrect, even if it affects the superstructure built on such "maintained" hypotheses. It is especially unreasonable to reject such an experimental approach because . . . the statistical theory which forbids the rejection of a "maintained" hypothesis is not fully satisfactory either in view of the difficulty of its application. *What is incorrect however, is to act as if the final hypothesis presented is the first one, whereas in fact it is the result of much experimentation. Since every econometric analysis is an essay in persuasion—just as is true for any other branch of science—the line of thought leading to the finally accepted result must be expounded.* [Theil (1961, pp. 206–207), emphasis added.]

The question Theil did not answer, but one which has begun to be answered in the recent literature (cited in footnote 3), is how the estimates are affected by alternative "lines of thought" or estimation strategies. While a review of this literature is beyond the scope of this comment, it is important to provide some general remarks on the overall implications of these efforts since they relate to all three of the papers on informal models.

The literature on specification searches has clearly identified the trade-offs involved in estimation. Each approach to data analysis does involve a trade-off between bias and precision (i.e., variance). Increases in the prior information imposed on the organization of sample information will increase the precision of that analysis. However, this gain is realized at the cost of increased bias should the prior information be incorrect. Since estimation strategies affect the manner in which prior information is selected and used, they can be discussed in terms of the trade-off between bias and variance which they imply. This conclusion is also true for the techniques organized under the broad heading of exploratory data analysis. It does not, however, imply that they should be dismissed as irrelevant. Rather it suggests that judgements as to the appropriate avenues for complementarity between exploratory data analysis and econometrics should be based on more realistic appraisals of the implications which these methods have for the quality (in terms of bias and precision) of the insights (i.e., the estimates) they provide with a given data set.

In this regard Mayer's example of modeling the demand for energy is an unfortunate one. It may lead the reader to the conclusion that economics has little to say about modeling the demand for energy and that one must rely on the data alone to "tell their story" to the skillful "masseur." There are nearly two decades of economic analysis to draw upon in understanding the household demand for energy. For the most part early studies were constrained by the data available not by theory. Indeed, most economic analysts recognize that, in the short run, the household demand for energy is a derived demand resulting from the utilization of a given set of appliances. For example, one can draw a direct parallel between the modeling of the household demand for natural gas and Fisher & Kaysen's (1962) early work on electricity. This parallel extends both to the relationship between the demand for natural gas and the utilization of the household's durable equipment (in Mayer's case space heating appliances) and to the nature of the price schedules. While many issues are not resolved in this area,[4] one would nonetheless be incorrect to conclude that economic models have little to contribute. In fact, on the basis of the simple analysis sketched below we might well conjecture that there would be little scope for a short-run price effect in a model of residential demand for gas.

More specifically, Mayer's empirical analysis relates to a group of relatively new townhouses of two distinct types. The data consist of observations recorded on each household's monthly use of natural gas over a five year period. Using the basic logic of the Fisher–Kaysen model and *no* data analysis one can readily observe that since the applicances in the townhouses (the space heaters) were new, any changes in the replacement cycle for them, even as a result of exceptionally large price increases, would not be likely to be observed during the initial five year period following their installation. Hence there is little possibility of measuring the effects of the long-run, and presumably more substantial, responses to price changes. Over the period studied the only avenue for a price effect is through the impact of the price on the utilization function for each appliance. Here, the potential for substitution of sweaters for increased room temperatures, insulation, and the like must be considered as among the potential responses households can make to reduce utilization rates with dramatic price increases. Unfortunately for Mayer's example the actual period in question was *not* one where dramatic price increases were observed. It is reasonable to expect that the majority of the observed changes in real price (average, marginal, or some other measure of the real cost corresponding to an alternative function of

[4] Taylor's (1975) review for electricity gives some useful perspectives on some of the issues involved in electricity demand.

the price schedule for natural gas facing each household) was due to inflation in the overall price level.[5] Thus we would expect little scope of any response *a priori*. Finally, the data do not permit the identification of the socio-economic characteristics of each household, such as income, family size, and ages of children. These may well be among the most significant determinants of variations in use of the households in the sample. Thus it would seem that a little economic analysis at the outset can contribute to the way in which an analyst would proceed to further data inspection with this particular data set.

2. Partial Least Squares (PLS)

The paper on the evaluation of informal models by Wold reviews the use of iterative methods for estimating a variety of different models in which there are latent variables. While the specific formulation of partial least squares depends on the structure under study, it can be considered as a general class of methods that calls for repeated application of ordinary least squares to a progressively altered set of variables in the estimation of both the parameters of a model and of the latent variables. These models are generalizations of the iterative methods developed by Wold and his colleagues for simultaneous equation models, for example, the fixed-point and iterative instrumental estimators.

As in the case of the estimators for simultaneous equation models, it is difficult to analyze the analytical properties of PLS. One suspects that they will be model specific. This conjecture is based on the recent analyses of similar estimators, such as the fixed point estimator under alternative modeling assumptions [see Maddala (1971) and Mitchell (1974) as examples]. Indeed, Wold acknowledges the potential for such problems in an earlier paper, noting that, "The problem of convergence [of PLS] has to be dealt with from case to case," [Wold (1974, p. 71), bracketed term inserted].

Therefore it seems reasonable to suggest the need for sampling studies, similar to those conducted by Mosbaek & Wold (1970) in the case of the fixed point estimator, in order to evaluate the PLS methods under a variety of conditions. This analysis is required before one can recommend the PLS procedure for widespread use in cases where the prior theory may offer little guidance as to the nature of the latent variables.

[5] Mayer's demand models, which include *both* a real price measure and what is designated as an inflation index, reflect this conclusion. One might well question the rationale for this formulation of a demand function.

3. Set Predictions and Econometric Models

Hildebrand, Laing, & Rosenthal have proposed that a general prediction logic developed to evaluate predictions and describe data in the analysis of samples in cross-classified format may be useful in some types of econometric analysis. Specifically, the authors suggest that those cases where a model does not lead to a single point prediction but to a range of values for the variable of interest, i.e., to a *set* prediction, may be especially suited to their prediction measures. The proposed measure of prediction success is the proportionate reduction in the error associated with those general models that account for the joint association between a dependent variable and one or more conditioning variables. This suggestion is an interesting general measure that parallels the square of the correlation coefficient for one variable in linear regression models.

The authors' proposals deserve serious consideration and should be more closely examined for a range of economic applications. However, they should not be evaluated in isolation. Indeed, we may wish to consider several econometric methods developed to meet the special requirements of models with truncated and multiple-valued discrete dependent variables. For example, Amemiya (1975) has recently proposed generalizations, which are based on the earlier work of Cox (1966), to logistic models to take account of those cases where more than two responses are possible. In the trichotomous unordered case such a structure could be represented as follows

$$P(Y_i \in S_1) = 1/\theta_i, \tag{1}$$

$$P(Y_i \in S_2) = \exp(\beta_1 X_i)/\theta_i, \tag{2}$$

$$P(Y_i \in S_3) = \exp(\beta_2 X_i)/\theta_i, \tag{3}$$

where S_1, S_2, S_3 represent a mutually exclusive and exhaustive set of values for Y_i, $\theta_i = 1 + \exp(\beta_1 X_i) + \exp(\beta_2 X_i)$, X_i is the set of determinants of the set predictions (a vector of variables for ith observation), and β_1, β_2 are conformably dimensioned parameter vectors for X_i.

There is no reason, in principle, why models such as this one could not be applied to the cases discussed by Hildebrand *et al.* These more formal specifications also have advantages. One of these is that they offer a more explicit means of incorporating information on the nature of the determinants of the types of set predictions, i.e., the outcomes S_1, S_2, or S_3, without the requirement that the relationships be strictly continuous in Y_i. Moreover, they are amenable to maximum likelihood estimation and, through the imposition of restrictions on β_1 and β_2, can represent a menu of possible associations between the prediction sets. Of course these models incorporate more prior information than is assumed in the methods of evaluation proposed by Hildebrand *et al.* Thus, here again, a comparison can be made which illustrates how the trade-off between bias and variance can enter the modeling and evaluation of informal models.

4. Summary

Each of the papers discussed here provides suggestions for the improvement in the methodologies used with informal models. As the introduction noted, these models are of growing importance for a wide variety of problems which often require the innovative application of existing methods as well as the development of new ones. In the process it is important not to lose sight of the fact that estimation should not be a mechanical process. If it is to be informative it is important to acquire as complete knowledge of the strengths and weaknesses of the data, the methods, and the practice by which these methods are applied as is feasible in each application.

REFERENCES

Amemiya, T. Qualitative response models. *Annals of Economic and Social Measurement*, 1975, 4 (September), 363–372.

Coelen, S. P., & McIntyre, R. J. An econometric model of pronatalist and abortion policies. *Journal of Political Economy*, 1978, **86** (December), 1077–1102.

Cox, D. R. Some procedures connected with the logistic qualitative response curve. In F. N. David (Ed.), *Research papers in statistics*. New York: Wiley, 1966. Pp. 55–71.

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

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

Hildebrand, D. K., Laing, J. D., & Rosenthal, H. Prediction analysis and econometric models, This volume.

Judge, G. C., Bock, M. E., & Yancey, T. A. Post data model evaluation. *Review of Economics and Statistics*, 1974, **56** (May), 245–254.

Lave, L. B., & Seskin, E. P. *Air pollution and human health*. Baltimore, Md.: Johns Hopkins Press, 1977.

Leamer, E. E. *Specification searches: Ad hoc inference with nonexperimental data*. New York: Wiley, 1978.

Maddala, G. S. Simultaneous estimation modes for large and medium sized econometric models. *Review of Economic Studies*, 1971, **38**, 435–445.

Mayer, L. S. The use of exploratory methods in economic analysis: Analyzing residential energy demand. This volume.

Mitchell, B. M. Fixed point estimation of econometric models. *Australian Economic Papers*, 1974, **13** (December), 250–266.

Mosbaek, E. J., & Wold, H. *Interdependent systems: Structure and estimation* Amsterdam: North-Holland Publ., 1970.

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.

Taylor, L . D. The demand for electricity: A survey. *Bell Journal of Economics*, 1975, **6** (Spring), 74–110.

Theil, H. *Economic forecasts and policy.* Amsterdam: North-Holland Publ., 1961.

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

Wallace, T. D. Pre-test estimation in regression: A survey. *American Journal of Agricultural Economics*, 1977, **59** (August), 431–443.

Wold, H. Causal flows with latent variables. *European Economic Review*, 1974, **5** (1), 67–86.

Wold, H. Model consideration and evaluation when theoretical knowledge is scarce: An example of the use of partial least squares. This volume.