This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Econometric Models of Cyclical Behavior, Vols. 1 and 2

Volume Author/Editor: Bert G. Hickman, ed.

Volume Publisher: UMI

Volume ISBN: 0-870-14232-1

Volume URL: http://www.nber.org/books/hick72-1

Publication Date: 1972

Chapter Title: A Note On Scientific Method In Forecasting

Chapter Author: V Lewis Bassie

Chapter URL: http://www.nber.org/chapters/c2793

Chapter pages in book: (p. 1211 - 1218)

A NOTE ON SCIENTIFIC METHOD IN FORECASTING

V LEWIS BASSIE · University of Illinois

THIS note is written as a discussion of forecasting methodology in general, rather than that of any particular models. I recognize that we are in a period of transition toward more mathematical, computer-based techniques of analysis. However, up to this point, there has been no breakthrough—either in theory, or in practical economic forecasting—as a result of econometric model-building, and any gains achieved in the last decade are certainly minimal.

What impresses me is how much closer together the judgmental and the econometric model-builders have come during this postwar period. At the 1951 Conference on Short-Term Forecasting, I was critical of econometrics on the ground that it was too inflexible for good forecasting. A more than adequate response to that criticism makes it no longer appropriate. But the enthusiasm with which the econometricians have gone about adapting their models to the practical task of matching a series of quarterly observations has gone beyond necessity in making concessions to the "flexible analysis" point of view.

Last year, I made an analysis of the 1966 OBE Model in terms of the relative contributions of exogenous and endogenous factors to a year-to-year forecast of gross national product. This analysis revealed that the forecast was largely predetermined by autonomous elements, trends, lagged dependent variables, and special adjustments. All government purchases, all exports, almost all nonresidential fixed investment, well over half of residential investment, and about half of consumption expenditures were derived exogenously. To understand the role of judgment in this: for *autonomous,* read *acceptance of somebody's projection;* for *trend,* substitute *noncausal catchall;* for *lagged dependent variable,* put *reliance on continuity;* special adjustment speaks for itself. With all that judgmental determination, the power of the analytical technique can make little difference in the forecast.

The great, guiding goal of the pioneers in econometrics was to make economic analysis scientific. The general import of the papers presented here reveals that in trying to be practical men who can meet the test of forecasting, their successors have resorted to various expedients whose main contribution lies in improving the fit to a limited series of observations. Much of what is being done is inconsistent with the scientific approach; and to that extent, model-building has tended to go off on the wrong track. However, this description of the current state of affairs probably reflects the model-builders' impatience and temporary frustration resulting from obstacles to getting the job done; so the final answer is yet to be given.

Since my undergraduate days, I have accepted the definition of scientific method in the social sciences to be the application of impartial intelligence to the solution of problems experienced in the real world. Here, as in the natural sciences, the process begins with the collection and ordering of facts to develop their implications, proceeds with the formulation of hypotheses to explain what produced the observed results, and finally goes on to test each hypothesis, not only against the known facts, but also against additional facts that can be collected to confirm or deny its validity. It seems to me that some comments are in order under each of these three headings.

THE FACTUAL CONTENT OF THE DATA USED

THE current fashion is to construct a model on the basis of quarterly data for a limited period of years in the 1950's and 1960's. Unfortunately, this ignores great masses of material available in the economic records; and even if the recent data are taken at face value, they are not well suited to the burden put upon them. All they can reflect is the pattern of interrelated movements that has been pervasive in these years of rising prosperity supported by Cold War tensions and by rapidly expanding Social Security and local government programs. The postwar period as a whole has been so consistent in its rarely interrupted uptrends that it is hard to conceive of this special collection of facts as being more than self-defining.

I recognize that quarterly data must be used to get the advantage of some variation that is smoothed out of the annual data. The latter are clearly not adequate for solutions by multivariate methods; they are too few in number and smooth all the series into an impossible mess of multicollinearity. But the shift to the quarterly data discloses only a small amount of independent variation; the autocorrelation in the quarterly data is often high when it is absent in the annual data. It is possible to go further in obtaining greater variation by shifting to monthly data. In a study at Illinois, we constructed a monthly gross national product series, but we found it of very little use for analysis because erratic elements were too important, especially in such series as the change in inventories and net foreign investment. Something of the same effect is retained in the quarterly data also, and it is inevitable that they should contain errors of adjustment as well as of observation. These are matters that urgently need attention.

A related question concerns the nature of a significant observation. It is evident that economic processes vary widely in the time spans through which they operate and that the durations which provide an adequate measure of the forces at work vary correspondingly. The time stratification that results in quarterly data is a compromise suitable to the analysis of some processes, say, the Keynesian multiplier, but can hardly be appropriate for others, such as the interacting cyclical forces that are characteristic of the processes involved in various kinds of capital formation. Distributed lags are, of course, important in many of these processes, and Franco Modigliani's analysis of new orders, unfilled orders, and expenditures is a masterly attempt to deal with this problem. There is hardly any comparable validity in cumulating lagged dependent variables over a number of quarters, as Ronald Cooper attests in presenting his own "auto-regressive schemes," even though improved estimates may be obtained in this way. The cyclical or other processes must be understood in order to specify the significant form of the data, and the way data are to be used should be an essential item of feedback for directing the work of the data compilers.

Even if it is assumed that we are adding to the statistical records of economic performance in an appropriate way, there is no assurance of early success in obtaining the information needed for determining appropriate model specification. For example, we are not in a much better position to specify a valid consumption function than we were two decades ago. The various series available for use as instruments have all been behaving so much alike that it cannot be definitely determined which are superior. Now we are in another period of disturbance whose duration we cannot readily estimate. This taxes the patience of all of us.

Furthermore it may be suspected that some of the things we have recently learned will prove to have only temporary validity. But it is possible, also, to go back before World War II in creating a factual basis for analysis. This will require pooling of data from different periods, with adjustments as necessary, but these tasks are not a diversion of scientific effort. Unless part of our efforts are devoted to improving the data over a long period of time, the wait for valid results will be prolonged still further.

THE HYPOTHESES UNDERLYING MODEL SPECIFICATION

MUCH of what has been done in selecting variables and specifying relationships departs from the basic goal of explaining observed facts as the result of causes. I heartily concur in the remarks of Robert Eisner and Charles Holt to the effect that analysis can be adequate only when it is logical and objective. There appears to be a tendency to look ahead toward the testing, that is, to look for instruments and equations that will provide good tracking over a limited period without regard to their theoretical underpinnings. This tendency stands in sharp contrast to the effort to establish scientific principles that will explain the past in such a way as to provide understanding of the present and of the emergent future we are seeking to predict.

It is not enough to say that this is *related* to that. Practically everything on the economic scene is related to other things through some indirect chain of interaction; specifying mere simultaneity of movement as a working relationship cannot be productive in the long run. Reliance on simultaneous movements is satisfactory as long as the variables continue to move together. But if it cannot be shown that they have to move together, the forecasts remain doubtful. Attaching gadgets to models for no reason other than that they happen to work for the time being is without scientific foundation. Cooper is aware that his own auto-regressive schemes can be expected to work only as long as the conditions prevailing in the period of fitting persist. I am impressed with his paper, and regret only that he was not able to take up, also, the evaluation of underlying theory, from which some even more important contributions might derive.

Some of the models incorporating mechanical constructs seem to ignore this point. It should be apparent that the use of trends, or growth rates, will sustain an upward bias through periods of recession and result in inadequate cyclical amplitude, as reported; and there is no good reason to think that some of the trend elements of the postwar period are really permanent. Again, the use of lagged dependent variables represents an application of continuity with only limited validity and tends to result in missing the turning points, as also reported.

Some of the discussion of periods of fitting sounds a bit outside the realm of reality. If shifting the period of fitting by one year—from 1953–64 to 1954–65—can make a significant difference in the model, then a shifting of conditions in any succeeding year can invalidate it. One long-standing rule is that the period of fitting should cover full cycles, because starting at the bottom of one cycle and ending at the top of another introduces an upward bias. Gaining the advantage of such a bias could, at best, have a short-term payoff. It is not at all surprising that some of the parameters in a complex model should be sensitive to data changes, whether the latter occur from shifts in periods of fitting or from other sources.

The scientific goal for a model is to represent the basic structure of the economy, with specified modifications for known changes in structure, such as tax legislation and the development of welfare institutions. Given that goal, there are no grounds for fitting to any short period and ignoring the longer perspective that would be obtained by extending the analysis backward through time, as well as forward with the accumulation of new experience. What we should abstract from are the special features of particular subperiods. In other words, a basic model should not try to fit everything. Some events or fluctuations should not be fitted because of their unique character as things that have happened once and may never happen again.

What this suggests is that the prevailing theories of stochastic variation are not very satisfactory. From a fully deterministic standpoint, the deviations are caused, too. Since it is impossible to analyze deterministically the full complexity of economic relationships and events, we assume that the best procedure is to isolate the dominant lines of causation and treat the remainder as determined by invariate probabilities. The rules of fitting then result in residuals and errors of estimate that appear to be randomly distributed. This remains the case where variables are stochastically determined, as in models incorporating Markov chains. But the assumption of stable probabilities is provisional, and the appearance of randomness in the residuals adds hardly anything at all. We cannot really say that the residuals are caused by "other" causes, because we do not know how much of the estimates themselves are caused by other causes not specified in the model. Where there is a choice between deterministic and stochastic approaches, the deterministic alternative should be selected, unless there are decisive reasons for not doing so.

The evidence of structural instability has continued to pile up. In 1951, I pointed out that few relationships can be considered invariant.[1] The operation of each is subject to a range of reaction that depends upon the prevailing conditions and the effects of other relationships whose influence under the circumstances may also be subdued or aggravated.

Knowledge of special situations tells us that many of the deviations are not caused by stochastic forces. In some cases that knowledge would be adequate for adjusting the data first and then readjusting to put the actual variation back in after the fitting. This procedure could well be superior to the use of dummy variables and would let the large recorded deviations of the period stand as reminders of the special events that accounted for them.

Dummy variables at their best have poor extrapolation qualities, and since the burden is thrown back upon the judgment of the analyst in any case, there is no harm in a direct estimate. The substitution of an indirect for a direct judgment does not in itself make an estimate more scientific, any more than the substitution of a mathematical symbol for a word makes the referent more objective. Extrapolation qualities

[1] *Short-Term Economic Forecasting,* Lawrence R. Klein, ed., Studies in Income and Wealth, Vol. 17. New York, NBER, 1955, pp. 31–32.

should, in general, be considered in specifying any variable or equation which is to be included in a forecasting model.

I believe most economists, as well as most businessmen and government officials today, seriously underestimate the instabilities in our economy. The fact that inflationary expectations dominate does not mean that there is, in fact, less danger on the downside. On the contrary, for whatever such attitudes may contribute to a booming economy, they make instability greater. I doubt that any model standardized on some subperiod of the last two decades of prosperity—and with elements of instability therefore reduced to a minimum—will be successful after conditions change.

THE ADEQUACY OF STATISTICAL TESTING

ANY testing must, of course, be carried out in terms of the facts, but statistical accuracy does not by itself ensure reliability in forecasting. The Conference has again encountered findings that confirm the long record of earlier studies which have shown how economic data can be matched by meaningless mechanical devices. In the past, the Foundation for the Study of Cycles was able to approximate the movements of the stock market over a period of eighty years by means of a trend line and eleven fixed cycles. One of my students made an excellent representation of quarterly real gross national product for the period 1953–66 by the formula

$$\frac{100\Delta G}{G_{-1}} = 1 + \frac{H - T}{2}$$

where $H + T = 10$, the heads and tails in a series of tosses of ten coins, with an ex post matching of the time scale. Other nonsense correlations are numerous. With so many ways to depict a given set of observations, both the underlying rationale and the longer perspective become important criteria for validity.

Testing the accuracy of projections one quarter ahead on an ex post basis is a trivial exercise; it would be absurd to construct an elaborate model for these short-range projections. The six-month ex post errors are still constrained, and estimates prepared on this basis should not be presented as an indication of results to be obtained in actual application. Still further out, any pretense of objectivity in such conceptions as "ex post with initial endogenous estimates" has no place in scientific analysis.

Furthermore, the whole ex post procedure of testing ignores an essential consideration in obtaining consistency in ex ante forecasts, namely, the need for reflexive adjustment of the autonomous variables. It is clear that as the situation changes, government programs and business investment plans will not remain unchanged, but will be revised in the light of what has developed. Unless the forecaster is willing to extend his hypotheses to cover such contingencies, his model will be subject to qualification in actual use.

Many participants in the Conference have pointed out that much remains to be done. In my view, problems needing particular stress are data extension and improvement, and a reorientation of specification procedures toward basic scientific methodology. Both of these should take precedence over data manipulation and success in approximating a limited set of estimates. Conformity to preconceptions—whether of fact, or of method—cannot be the route to scientific progress. Both causality, as a rationalizing concept, and sophisticated statistical technique, as a means of establishing relationships, must be stages in a more complete process of open-minded intelligence by which the solution of problems is achieved.