

Selected Papers • No. 15

Forecasting
Business
Conditions:
A Critical View

By VICTOR ZARNOWITZ



GRADUATE SCHOOL OF BUSINESS
UNIVERSITY OF CHICAGO

VICTOR ZARNOWITZ is Professor of Finance in the Graduate School of Business of the University of Chicago. His main research interests are in the areas of macroeconomic theory, business fluctuations, and time-series analysis. He received his Ph.D., *summa cum laude*, from the University of Heidelberg in 1951. A native of Poland, he acquired American citizenship in 1957. He has been a member of the research staff of the National Bureau of Economic Research for the past dozen years. In 1953-54, he held a Social Science Research Council postdoctoral fellowship at Harvard University; lectured at Columbia University between 1956 and 1959; and was awarded a Ford Foundation faculty research fellowship in 1963-64, to pursue studies of the variability of investment demand. He joined the faculty of the Graduate School of Business in 1959.

. Mr. Zarnowitz is the author of monographs and papers on the theory of income distribution, business cycle indicators, and the cyclical behavior of manufacturers' orders and prices. Recently, he has been directing a study of the accuracy of short-term economic forecasts for the National Bureau of Economic Research. This Selected Paper is based in part upon a talk he delivered at the 6th Annual Meeting of the National Association of Business Economists in Philadelphia on September 29, 1964.

Forecasting

Business Conditions:

A Critical View

SHORT-TERM economic forecasting, as widely practiced today, is largely an art or game ruled at best by experienced and disciplined judgment, at worst by sheer luck. It probably will always contain much of these elements, but a trend is to be expected toward increased application of scientific methods of evaluating the evidence and drawing inferences from it.

In recent years, forecasts of the nation's economic fortunes have become much more abundant and ambitious than ever before. It is now possible to assemble a fair-sized collection of continuous forecasts from well-reputed sources for several major economic aggregates and indexes, as we have done in a study currently underway at the National Bureau of Economic Research. As one tries to extend the record back to the early postwar and the prewar years, sources quickly dry out.

That forecasts grew in boldness as well as in quantity can be seen in the fact that many are now expressed in specific numbers. Vague, hedged, or purely qualitative predictions of "what's ahead for business" are still quite common, but they no longer dominate. Also, attempts are increasingly made to predict the course of the economy over a sequence of short periods-say, the four quarters of the year ahead-and this represents a particularly ambitious, "dynamic" type of forecasting.

These developments reflect increased demand for forecasts of economic conditions. Business management clearly has a very large share in that demand, and its preference is for unconditional, specific, numerical predictions. However, the demand for forecasts is diversi-

fied as well as large. For example, the forecasts required by government policy makers (another large source of demand) differ from those sought by the business and financial community in that forecasts conditional upon alternative policy courses are precisely what is required by the former.

Changes contributing to the growth and specialization of forecasters' output occurred also on the supply side of the market for new "economic intelligence." The amazingly rapid development of electronic computer technology accelerated greatly the rate at which economic data (the raw materials for the forecaster) are compiled and processed. It also had some more direct effects-without the computer, the large-scale econometric models could not have been produced, hence output of forecasts of the econometric variety would have been limited. However, work with such models is still essentially in the domain of academic economists. The great majority of forecasts are produced by business economists, who have thus far appeared to make very little use of formal econometric models.

Whatever services the forecasts are expected to render to the user, they vary, and are not easily defined by an outside observer. However, the usefulness of forecasts is surely in the first place a function of their accuracy.

With the growth of public interest in the expanding activity of economic forecasting, there is increasing need for objective and comprehensive evaluations of the forecasts. This need is as yet largely unsatisfied. It is surprising to note how little systematic testing has been done in this area-despite the widespread use of business activity forecasts, their costs, the potential rewards of good predictions and penalties of bad ones, and the consequent importance of the ability to discriminate among the available sources and methods.

In attempting to survey the field we must

ask ourselves: How is the accuracy of economic forecasts to be assessed? How useful are such assessments as can be made, in terms of a quantitative description of forecasting errors—their magnitudes, types, and structure? What inferences can be drawn about the dependability and usefulness of the forecasts? What about the feasibility of improvements? Though research on these problems has lagged badly, we have recently begun to make some real headway in attacking these questions.

This is not the place to review the literature on the subject, but one recent study is so important that it must be mentioned: The work of Henri Theil on the methodology of forecast evaluation and the accuracy analysis of certain European forecasts both of the business survey and the econometric-model variety.¹ And the current National Bureau study of U.S. materials, in which I am engaged, already has produced some interesting findings about the accuracy and other characteristics of aggregative short-term forecasts. Let me now turn to some central issues in economic forecasting, as revealed by these recent explorations.

The Hazards of Economic Forecasting

That economic forecasting is a hazardous art is common knowledge. The precise reasons for this are not so well understood.

It is of some help here to consider the typical economic time series to be predicted as a composite of four factors: trend, cyclical, seasonal, and purely “irregular” or random movements.

Trend fittings and projections are often relatively successful in application to long-term forecasts, but in many cases they can play only a subordinate role in the short-run context. (However, for some series trends are important even over short periods, and will often be well approximated by simple methods, which

¹ Henri Theil, *Economic Forecasts and Policy*, North-Holland Publishing Company, Amsterdam, 1958.

should facilitate the forecaster's task considerably.)

Strictly periodic, repetitive fluctuations should also be rather easy to predict: stable seasonal movements would often be more or less of this type, but forecasters usually work around them, trying to forecast the "seasonally adjusted" series.

This leaves the cyclical and irregular components as the main sources of trouble for the short-term forecaster. Looking forward, it is anything but easy to distinguish the cyclical from the random element in the movement of an economic time series, though retrospectively it is usually possible to do so with fair results.² The forecasting errors that are directly traceable to very short random movements must really be accepted as unavoidable. The forecaster can hardly be expected to predict an event generally regarded as unforeseeable such as an outbreak of a war (e.g., Korea, 1950) or a strike started without advance warning. Though such "shocks" cannot themselves be predicted by the techniques of economics, their more significant effects on the economy are, of course, the proper concern of the forecaster.

The requirement of a good forecast is that it predict well the systematic movements, trends and cycles—not that it predict perfectly the actual values of the variables concerned (it could not do that, except by accident, for economic series—which, as a rule, contain random elements). And of the "systematic" movements it is the cyclical fluctuations, not the longer trends, that produce the greatest difficulties in short-term forecasting. These fluctuations are recurrent but nonperiodic; they vary greatly in duration and amplitude; calling them cyclical

2 The practice used most frequently for this purpose is to pass a moving average of intermediate length through the seasonally adjusted series and get the deviations of the smoothed from the unsmoothed values as estimates of the irregular component (which can then be tested for their randomness properties).

should convey neither more nor less than that they reflect mainly the participation of the given economic factor in "the business cycle."

I shall illustrate later the importance of the business cycle as a source of forecasting errors. Meanwhile, some related trouble-making factors must be noted. One is the lack of accurate information about the conditions prevailing at the time the forecast is made. The initial level or base from which the predicted change is measured must itself be predicted; and although they are estimated at a close range, the base figures often contain significant errors. For example, in predicting the level of GNP next year, errors made in estimating the base often contribute as much as 30-40 per cent to the total forecast error.

Better estimates of current position could improve substantially the forecasts themselves. Moreover, such common targets of forecasters as the nation's aggregate output are exceedingly difficult to measure or even to define. The forecasts also frequently involve factors of presumed importance which are very elusive, such as the state of "business confidence." Where measurement is difficult, and estimates can have substantial errors, prediction seems particularly hazardous.

Size and Direction of Errors

Let us now take a closer look at the accuracy of forecasts of business conditions.

The errors of the annual forecasts of the gross national product (GNP) in current dollars (both over- and underestimates) averaged about 7-11 billion dollars in the years 1953-1965. The change in GNP during this period averaged about \$22 billion per year.³ The errors, then, tended to be approximately one-third to one-half the size of errors that would be produced by a "naive model" which as-

3 Computed without regard to sign.

sumed the predicted year's level of GNP to be the same as the previous year's.

Compared with average levels of GNP, these errors would amount to no more than $1\frac{1}{2}$ - $2\frac{1}{2}$ per cent. But short-period changes in GNP are generally small relative to already attained levels (their order of magnitude here is 4-5 per cent). The recent recessions were all short and relatively mild. A margin of plus-minus two per cent of GNP can mark the difference between a "good" and a "bad" year.

Forecasts of average levels of economic activity in the coming year (expressed in terms of GNP and its major components and the industrial production index) are in general more accurate than mere mechanical extrapolations of the past. The forecasts examined proved superior not only to the simplest "naive model" extrapolations of the last known level or change, but also to some much more demanding standards of trend extrapolations and autoregression models.⁴ In "beating" the predictions produced on the computer by weighting, averaging, and extrapolating past values of the 'given series, the annual and shorter forecasts by economists scored what must be regarded as a significant success—even though the best of the mechanical yardsticks against which they were measured leave something to be desired, and occasionally the margins of success have been slight.

Business forecasters have been called cautious or conservative because of their tendency to underestimate changes. The data analyzed in the current National Bureau study confirm this observation emphatically. Underestimation applies to both increases and decreases, and is evident in forecasts relating to different variables. Underestimation of increases usually results in underestimation of the ensuing levels. Underestimation of decreases, analogously,

4 Statistical models of the relationship between present value and past values of a given economic series.

tends to result in over-estimation of levels. In series with upward trends, such as GNP or industrial production, increases are more frequent than decreases so that the levels are understated most of the time.

To the extent that underestimation reflected merely a smoothing out of the random component of the actual values in the forecasting process, it could not be objected to as a type of systematic error. There is ample evidence, however, that the observable tendency goes far beyond that and constitutes indeed a true bias. The average changes in actual values generally exceed those in predicted values, whether taken with or without regard to sign.

The tendency to underestimate changes is on the whole stronger in simple mechanical extrapolations than in the forecasts proper over short spans of time. But forecasts and extrapolations probably have much in common on this point. The primary dependence on the data of the recent situation itself can be a basic source of the bias. Forecasters necessarily rely on the stability of some relationships observed in the past, but these in fact are undergoing changes. As a result, stability is exaggerated; that is, change is understated.

We find that forecasts of rather good quality often have been made for the very near future -the next quarter or six months. (The average annual forecasts can be viewed as having mean spans of little more than six months, too.) However, with further extension of their reach into the future, short-term forecasts deteriorate rapidly.

A few examples will be sufficient to demonstrate the regularity and pervasiveness of this relation. In a semi-annual forecast of GNP for 1955-63, representing an average of a fairly large group of individual predictions, the mean absolute errors of the relative change

are: for a six-month span 1.5 per cent, for a twelve-month span 2.6 per cent. In a quarterly forecast by the staff of a large company for the same years, the mean absolute errors computed analogously in per cent are: for one quarter 1.1, for two quarters 1.8, for three, four, five, and six quarters 2.5, 2.9, 3.4, and 3.8, respectively. Again, in an semi-annual forecast of the Federal Reserve index of industrial production for 1947-63, the mean absolute errors are 2.8 per cent for six, 5.8 per cent for twelve, and 9.5 per cent for eighteen months!⁵

Why should the accuracy of short-term forecasts be a sharply decreasing function of the span of the forecast? Let us think of these forecasts as consisting of any or all of the following ingredients:

1. Extrapolation, of some kind, of the past behavior of the given series.
2. Relation of the series to be predicted to known or estimated values of some other variables.
3. Any other external information considered relevant, e.g., a survey of investment intentions or a government budget estimate.
4. The judgment of the forecaster.

Now it can be argued that each of these potential sources of the forecast is subject to a deterioration with the lengthening of the predictive span.

This is clear for the first ingredient: for example, a prediction that the level of industrial production in April will be the same as in March is likely to be less in error than the prediction that the level next September is going to equal that of March. What is here illus-

5 The last figure refers to the period 1947-55 (the strictly corresponding averages for the six- and twelve-month forecasts covering only those earlier years are 2.8 and 7.9 per cent, respectively).

trated on the simplest case is less obvious but still basically true for more sophisticated extrapolative or autoregressive models.

Informed judgments and estimates will probably also serve best over a #relatively short time range. The forecasting relations between time series involve lags of various lengths, but typically from some point on the relations weaken as the lags are increased. In the case of the so-called "leading indicators," which tend to precede the turning points in general business activity, several factors combine to reduce their effective forecasting lead. The earliest presumed signals of reversal must usually be confirmed by subsequent behavior of the same and other series. The evidence of a group of indicators is more reliable than that of individual members. Brief erratic variations often obscure movements of cyclical significance. Smoothing helps to bring out the major movements and turns in the indicators, and to reduce the number of false warnings, but it also cuts down the length of the effective forecasting lead.

While the forecasts with short spans are generally superior to extrapolations, those with longer spans (say of 12 to 18 months) are often worse than the more sophisticated types of extrapolation. For example, several of the recent forecasts of GNP came out poorer than the results of autoregression methods or simple trend estimates (such as projections of the average historical change in the series). Significantly, the failures included forecasts of GNP for the end of the next year, but not those covering the year as a whole (including its earlier parts), nor shorter-span forecasts from the same sources.

Forecasting and the Business Cycle

Cyclical movements are persistently undervalued by most forecasters. When forecast errors are averaged separately for different stages of the business cycle, it turns out that the levels

of major aggregates (GNP, industrial production) are understated most in the first year of an expansion. Later in the expansion, when these aggregates usually grow slower, their levels are underestimated much less. (Occasionally, they are overestimated, as in the unexpected retardation of 1962.) In contractions, forecasts as a rule exceed the actual levels, either because the downturn is missed or because the decline turns out to be larger than predicted. Such errors can be observed in forecasts with different spans, except for the longest ones (exceeding one year) where the errors are very large throughout and the differences within the cycle are statistically not significant.

Failures to recognize the turning point constitute another important category of "cyclical" forecasting errors. To appraise these errors, one must ask two questions: How often do turning points occur which have not been predicted? How often do predicted turns actually occur?

In annual forecasts of aggregates which tend to grow most of the time, "false signals" of turning points are understandably infrequent. Few reversals of direction will be here foreseen from one year to another, but rather increases will as a rule be expected. In forecasts that relate to shorter intervals and are issued more frequently, however, the false warnings are likely to be more troublesome. Given the span and frequency of the forecasts, errors of this kind will occur more often for variables with weaker trends but stronger cyclical and irregular movements than for the smoothly growing series.

Forecasters often failed to predict 50 per cent or more of the turning points that did occur. The proportion of these errors does not seem to depend systematically on the length of the predictive span. The hit-miss record of, say, fifty-fifty may appear worse than it actually is. One reason for this is that the record includes

some errors that are largely excusable. This clearly applies to the errors connected with the outbreak of the Korean War and its early economic consequences, an event described earlier as an externally caused "shock." Another large concentration of errors occurred in 1947-48, when an early postwar depression was widely anticipated. This was a grave misjudgment of the situation which cannot be exculpated just because it was very common, but one must also consider that the disruption of economic relationships caused by the war made the early postwar forecasts particularly vulnerable.

Furthermore, it is important to recognize that the ability to predict correctly at least some of the turning points, which the forecasts reviewed demonstrably have, is an advantage over the extrapolations, which in general cannot signalize turns at all. (The turns in extrapolations will as a rule lag those in the actual values; the strength of a good projection lies almost entirely in that it may predict well the longer-term trends.) An accuracy score of 40-60 per cent in predicting reversals, such as is found for many short-term aggregative forecasts, may be far from good but it is surely much better than zero. It is true that forecasts may signal many false or "extra" turns, which extrapolations could avoid. But this disadvantage may be outweighed by the advantage of the correct turning-point predictions, and there is evidence that it frequently is.

It appears that the forecasters have on the whole a better record in predicting upturns than in predicting downturns. In a contraction of the recent postwar variety, it should indeed be reasonable to start watching out for an upturn in, say, the third quarter of the movement, as pressures for an effective counterrecessionary action will have mounted and forces working for a recovery will have gained strength by that time. An analogous argument could be made for an expansion, where one

would use years instead of quarters, but here it may seem considerably less convincing.

One reason is that, in recent times, business expansions have varied much more in duration than contractions.⁶ Another related factor is economic policies. These are used to cut short the recessions they were unable to prevent, but their aim in expansions is to prolong the duration of the movement. Their effect, here as elsewhere, is of course not assured, and they can even prove perverse, and occasionally do. But forecasters who assumed that a recession will not be permitted to last beyond at most a year would not have been wrong in the recent times. On the other hand, the success of the policies to steady and lengthen the expansions was probably more difficult to gauge.

Moreover, the upward trend forces in the economy work in the same direction of lengthening expansions and shortening contractions, and this is indeed a generally powerful, and often decisive factor. The growth forces are not always equally intensive, however; when they temporarily slacken, the expansion will be weaker and shorter.

As an important qualification of this argument, it must be noted that a prompt recognition of a cyclical trough is conditioned upon the prior recognition of the peak; if the latter is delayed, the former may be frustrated. This brings up the subject of the aids to timely recognition of the turning points, that is, the series used as statistical indicators of business recessions and revivals.

At a cyclical reversal in business activity, the turns in "coincident aggregates"-such as pro-

^a The average duration of contractions was 10 months, with a standard deviation of 2 months; the corresponding figures for expansions are 35 and 8 months. Contractions ranged from 9 to 13 months; expansions from 25 to 45 months (this covers the period since 1948 but does not include the current expansion, which has already lasted four full years, having started in February, 1961).

duction, employment, and income-occur at approximately the same time, but there are some series that reach their turns earlier and others that tend to lag. Some indicators would be expected to lead others and do so: for example, new orders predict production of durable goods; housing starts predict residential construction; investment commitments (appropriations, contracts) predict plant and equipment expenditures. Since such indicators anticipate the movements of the coincident aggregates, their tendency to turn ahead of the general business recessions and revivals has a strong logical as well as empirical foundation. Their early timing provides the forecaster with an advantage for which there is no good substitute. They can, together with other anticipatory data (e.g., surveys of investment intentions) help rather efficiently at least to reduce the lag in recognizing the cyclical position of the economy. Thus a business recession may be identified at about the time of its occurrence or shortly thereafter, which is no mean achievement, considering the necessity to compensate for the delays in the collection and processing of the data and the fact that historically such events had a demonstrably long recognition lag.

We know already that at times better results yet are achieved: a turning point is predicted correctly ahead of the event. Evidence of early indicators and other anticipatory data may help produce such forecasts over short spans of time. However, good judgment must probably be given a large share of the credit (as poor judgment must be blamed on other occasions for unduly long lags of recognition). The same factor (if not simply the forecaster's luck) would certainly have to be credited for any successful turning-point predictions with longer spans, where the indicators could not have provided a sufficiently early signal. This is so because it is very difficult to find series

which would have effective leads in relation to the early cyclical indicators themselves.

Observations

Far-reaching changes in the structure and functioning of the economy (involving both its public and its private sector and their interaction) have made business fluctuations milder in the postwar period than they used to be in earlier times. The interest of academic economists in business cycles as a subject of systematic research has correspondingly weakened. However, despite important changes in some of their aspects, business cycles continued to be a recurring feature of the American economy.

It appears that economists have of late underestimated the importance of cyclical processes. An economic theorist may choose his subjects or view them in such a way as to avoid the troublesome business cycle problems, but a practical analyst and forecaster of business conditions cannot very well do so. However, he too tends to underrate the cyclical movements of the economy, as attested by the types of forecast errors I have discussed. There may be a useful lesson in the fact that the recent business fluctuations, even though mild, were not quite as minor as many had apparently anticipated.

Forecasts of business conditions in the near future are, in large part, considerably more accurate than several types of extrapolation, which is encouraging. It would be important to know how much of the success of a forecaster was due to his method or "model" and how much to his pure "judgment." In general, however, little can be said about this, since little is known about the business forecasters' assumptions and techniques. From some comparisons between business condition forecasts and more formal forecasts based on statistically estimated economic relationships, it appears

that judgment often does make a net positive contribution to the predictive performance. Yet it is also important to realize that it is the improvement in that is necessary if forecasting is to become more dependable; superior individual judgment is not replicable and not readily communicable to others. The best examples of the forecasting “art” may well surpass the results of a “scientific” forecasting from explicit models—the patterns observed in the past. But to the extent progress in this area is possible, it must nevertheless depend mainly on the development of the scientific rather than the artistic component of forecasting.