

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

Volume Title: An Appraisal of Short-Term Economic Forecasts

Volume Author/Editor: Victor Zarnowitz

Volume Publisher: UMI

Volume ISBN: 0-87014-445-6

Volume URL: <http://www.nber.org/books/zarn67-1>

Publication Date: 1967

Chapter Title: Introduction and Summary

Chapter Author: Victor Zarnowitz

Chapter URL: <http://www.nber.org/chapters/c0935>

Chapter pages in book: (p. 1 - 8)

## INTRODUCTION AND SUMMARY

THIS report is one of the products of the study of short-term economic forecasting, a current project of the National Bureau of Economic Research. The primary purpose of the project is to assess the accuracy of short-term forecasts of aggregate economic activity in the United States. The materials compiled and analyzed are authentic (*ex ante*) forecasts of the nation's economic fortunes for the near future—the next year or two or several shorter periods. The forecasts cover comprehensive measures of the value or volume of output for the economy as a whole and its main sectors: gross national product (GNP), its major components, and industrial production. Eventually, other variables will be added, such as personal income, employment, unemployment, and the main price indexes. The number of forecasts already collected is large and varied; when completed, the compilation should be a representative one and certainly far richer than any other such collection.

Our materials include both published and unpublished forecasts. Some of the latter were made available to the National Bureau on a confidential basis. But even for published forecasts, a decision was made at the outset of this study not to disclose the sources of the forecasts evaluated. Accordingly, none of the forecasts considered here is identified as to source. A word of explanation of this position is in order.

Forecasters draw to a large extent upon the same "raw materials," that is, on information that is widely accessible, and they variously influence one another. Few if any of them can be regarded as "independent producers," though some meet this description better than others. However, forecasters are rivals or competitors as well as co-

operators. Any statement bearing on the relative quality of a forecaster's product could be used in this competition. The National Bureau did not undertake the present study with the intention of supplying such information. Its purpose is scientific, and the decision on disclosure was taken just so that purpose would remain paramount.

Economic forecasts have several aspects or dimensions, which can be usefully measured and weighed in more than one way. Appraisals and comparisons of forecasts may yield significantly different results depending on how this is done. A sound evaluation requires that several measures be examined critically. It is not possible, then, to make the procedure very simple, lest the results be incomplete and therefore misleading. Intrinsic to the process of deriving appropriate measures of forecasting performance are certain checks and balances which help to make the appraisal objective and fair. But, by the same token, it is easy to distort the appraisal, simply by quoting selectively or out of context. In these circumstances the need to guard against possible misuse of the analysis is impelling. Moreover, the possibility of error in the analysis itself cannot be entirely eliminated.

For these reasons, the sources of the forecasts reviewed are not identified. We do not believe that this reduces the scientific value of the findings. Since we describe fully the methods of measuring forecasting accuracy that we have used, the same methods can readily be applied by others for purposes of comparison or verification.

The analysis of predictive accuracy yields a description of forecasting errors—their magnitude, type, and structure. The problem is to evaluate the errors in such a way as to make meaningful inferences about the dependability and usefulness of the forecasts.

#### PLAN OF STUDY

This study is divided into eight chapters. The first gives a short description of the nature and sources of the forecasts collected. The second presents some basic data and introduces several subjects that are treated in detail later, such as the measurement of forecast errors and comparisons between forecasters, forecasts, and extrapolations. It illustrates two types of informative arrangement of the data.

Chapter 3 compares the predicted and actual changes in each successive period and reviews the correlations among them and among the errors for different forecasters. It considers whether large errors are par-

ticularly frequent in some periods and rare in others, whether underestimates are concentrated in some periods and overestimates in others. It discusses how the predictive errors are affected by the cyclical characteristics of the forecast period.

Chapter 4 presents the over-all record of predictive accuracy for annual forecasts of GNP, its major expenditure components, and industrial production. It shows how errors in predicting the future were affected by errors in estimating the present, to what extent errors were systematic rather than random, how often the forecasters underestimated or overestimated the actual changes, and what was the forecasters' record on turning points. It shows which of the variables presented the forecasters with particular difficulties and considers the reasons. It also discusses the effects of aggregation over the GNP expenditure sectors.

Chapter 5 examines the characteristics of multiperiod forecasts, which try to predict a sequence of values for an economic variable. It analyzes the average relations between the time span of forecasts and their accuracy, and also the implicit marginal relations. Types of error in multiperiod forecasts (under- and overestimates, directional and turning-point errors) are also discussed.

Chapter 6 presents, for the variables concerned, extrapolations of the last-known levels and changes, of average historical changes (trends), and of the estimated relations between the present and the past values of the given series. It compares macroeconomic forecasts with these mechanical extrapolations and shows the requirements of the different extrapolative models. The relation between these comparisons and the length of the time span covered by the predictions is also analyzed.

Chapter 7 compares the merits and shortcomings of certain types of forecasts. It considers the possible advantages of sectoral forecasts by experts and the effects of aggregating forecasts by individuals into an average group forecast. It examines whether meaningful comparisons can be made between the records of different forecasters, whether some have been significantly better than others. It discusses the consistency of the forecasters' record over time.

The final chapter reports briefly on the progress made thus far in other areas covered by the National Bureau's study of economic forecasting and presents an outline for further research.

I have tried to make this report as simple and nontechnical as possible, but may not have succeeded very well in that endeavor. The reason, apart from any inadequacies of exposition, is that evaluation of forecasts is a relatively complex task which can hardly be done without some statistical tools and a modicum of technical language. As a rule, however, the passages that seemed more difficult have been relegated to footnotes which need not detain the reader who is uninterested in, or unprepared for, the complications or detail involved. There are only a few sections in which the text itself contains such passages, notably that on the bias in forecasting (in Chapter 4) and those on the extrapolative models (in Chapter 6).

The reader who wishes only to acquaint himself with the main results of this study may find it sufficient to limit his attention to the "conclusions" in the next section and the following parts of the text: Chapters 1-3; Chapter 4, first three sections; Chapter 5, first section; and Chapters 7 and 8.

#### CONCLUSIONS

This section lists our main findings to date. Some additional forecast data have been assembled but not yet processed, and research on some of the points mentioned is still under way. The conclusions that follow, therefore, are in part still tentative and provisional.

1. Records of forecasts of gross national product for the year ahead, covering the period between 1953 and 1963, show an average error, without regard to sign, of \$10 billion. Of this, \$1.8 billion is accounted for by errors in estimating the level of GNP at the time the forecast was made, and \$8.2 billion by errors in predicting the changes in GNP. These errors are computed by comparing forecasts with early recorded estimates. If current, revised GNP figures are used, considerably larger errors are obtained. This is so because in most years forecasts tended to underestimate the preliminary figures, while revisions tended to raise the levels of GNP.

2. The mean absolute error of about \$10 billion amounts to no more than 2 per cent of the average level of GNP in that period, but to a much larger fraction—approximately 40 per cent—of the average year-to-year change in GNP. Since the forecasting of changes is the primary objective of short-term forecasting, the latter percentage is the more significant measure of the degree of success or failure.

3. The mean absolute error of forecasts of industrial production for the same years was about 4 index points on the base 1947-49 = 100. This is about 2.7 per cent of the average level and 47 per cent of the average annual change in the production index.

4. The over-all averages conceal a considerable amount of dispersion among the average errors of different forecasters. These errors range from \$7 to \$14 billion for GNP forecasts between 1953 and 1963, and from 3 to 5 index points for industrial production forecasts. These figures still underestimate dispersion, inasmuch as the forecasts include some group averages which are themselves summaries of divergent individual predictions.

5. The predictions under study represent the product of several hundred forecasters. They were made by economists associated with industrial firms, banks and other financial companies, business publications, government agencies, and universities. Substantial differences between summary measures of error indicate that predictions made by individual economists or small teams for business enterprises are better than those produced by large groups or polls. This is so even though the average forecast for a group is in the long run typically more accurate than most of the forecasts of the individual members in this group because of compensating errors among the member forecasters.

6. There is also some indication that reliance on experts in particular sectors may yield better forecasts. However, it is extremely difficult to establish significant differences among the different types and sources of forecasts. The reasons are in part technical, such as the fact that forecasts for a given year are made at different dates and the late forecasts have an advantage over the earlier ones. Also, forecasts rank very differently in accuracy from one year to another.

7. Forecasts of comprehensive economic aggregates were in general more accurate in the 1953-63 period than in the early postwar period. However, there were certain special difficulties attached to the early postwar forecasts (notably the disturbances caused by World War II and later, in 1950-51, the outbreak of the Korean War and its early consequences). There is no evidence that forecasters' performance improved steadily over the period covered by the data.

8. Most forecasts underestimate the growth of the economy as measured by GNP. The underestimate is typically largest for the beginning

of a recovery from a business recession (when growth rates tend to be particularly high). Declines are less frequently underestimated than increases. Changes in series which fluctuate more and have grown less vigorously (e.g., gross private domestic investment) have been overestimated as often as underestimated.

9. Apart from the early postwar period, few significant errors were made in end-of-year forecasts on the direction of annual changes in comprehensive economic aggregates. The timing of recent business cycle downturns was early enough to make the presence of the recession widely known by the end of the peak years (1953, 1957, 1960). This, plus the presumption that the contractions would be short, made the task of predicting annual changes relatively easy.

10. Annual forecasts of GNP and industrial production are, on the whole, more accurate than any simple extrapolation of the preceding year's level or change. Even the more refined and effective kinds of extrapolations, based on relations between present and past values of the series or on recent average rates of change, are inferior to the economists' forecasts for these variables.

11. Forecasts of GNP and the production index for one to three quarters, like the annual ones, are typically better than all types of extrapolation. However, accuracy diminishes steadily as the forecast span increases. Forecasts for four quarters or more ahead are generally not superior to extrapolations of the recent trend (measured simply by the average rate of change). The record of year-to-year forecasts does not imply any greater accuracy than this because such forecasts are generally made late in the preceding calendar year and a good record in the first two quarters will produce a moderately good record for the year as a whole.

12. Marginal errors of multiperiod forecasts do not increase systematically, that is, average errors do not increase faster than the extension of the forecast span. In fact, decreases as well as increases in the marginal errors are observed, and no strong systematic tendency emerges when an adjustment is made for the fact that the recent past and present must in part also be predicted because of the lag of information. This evidence is consistent with the idea that projection of a certain rate of growth over a sequence of short intervals has been one of the basic devices in the construction of the multiperiod forecasts.

13. The multiperiod forecasts, which include predictions for several

short intervals, are more relevant for an appraisal of turning-point errors than the annual forecasts, and they present a different picture of such errors. The results here are, on the whole, negative: the record of the numerical forecasts of GNP (like that of qualitative turning-point forecasts) does not indicate an ability to forecast the turn several months ahead. Not only were actual turns missed but also turns were predicted that did not occur. Most turning-point errors were associated with declines in the given series. This is not surprising since downturns were more difficult to predict than upturns in the postwar period.

14. Forecasts for GNP as a whole are typically much better than those for most of the component types of expenditure. This reflects in part cancellation of errors in the summation by sectors (similar to the cancellation of individual errors of opposite sign in deriving group forecasts, noted under point 5 above). Another probable reason is that some methods of forecasting, such as the use of business cycle indicators or money supply, are concerned directly with measures of aggregate economic activity rather than with any component expenditures or sectors of the economy, and hence may yield better forecasts for total GNP.

15. The errors in forecasts of percentage changes in personal consumption expenditures are much smaller than those in forecasts of gross private domestic investment. Errors in predicting government spending are of intermediate size. The greatest *need* for improvement is in forecasts of construction (particularly residential), changes in inventories, and net foreign investment. Even though these series show relatively weak trends and strong cyclical and irregular movements, and hence do not lend themselves to effective extrapolations by any simple means, their forecasts have often proved to be little better or even worse than mere extrapolations.

16. The greatest *scope* for improvement, however, probably exists in the forecasts of consumption, including those for nondurable goods and services. This is because these are smoothly growing series which would have been predicted very well in recent years by simple trend extrapolations. The average errors of consumption forecasts have typically been larger than those of such extrapolations.

17. Aggregation of short-term expectations or plans of business concerns about their outlays on plant and equipment, as developed in



periodic intentions surveys, results in better predictions of total business capital expenditures than those made independently for the entire economy. This can be inferred from comparisons between the investment forecasts in our collection which are made before and after the McGraw-Hill Survey of Investment Intentions, and also from comparisons involving the Commerce-SEC investment anticipations data.

18. Better utilization of the historical content of the series could lead to a significant improvement of the forecasts. It would seem desirable that, at some stage, trend projections should be incorporated in the forecasting process since our study shows that many forecasts looking ahead four quarters or more are inferior to simple trend extrapolations.

19. It may be possible to achieve further gains by improving the shortest forecasts. Experiments show that knowledge of the next two quarters combined with even the simplest projections for the further future would lead to annual and longer forecasts that are better than those actually on record. This suggests that forecasters may do well to concentrate primarily on the two or at most the three nearest quarters: if they improved these short forecasts (which can be much better than mere extrapolations), they would also be able to produce better longer forecasts by means of projections over the more distant periods.

20. Improvements in the record-keeping practices of forecasters are badly needed. The records should not only include the forecast but also the estimated present position (errors in the latter are, as a rule, substantial and their measurability is important). The methods used to arrive at the forecasts should be specified. Such records would facilitate future appraisals, reveal limitations of forecasts, and might suggest improvements in techniques.