

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

Volume Title: Forecasting and Recognizing Business Cycle Turning Points

Volume Author/Editor: Rendigs Fels and C. Elton Hinshaw

Volume Publisher: NBER

Volume ISBN: 0-870-14479-0

Volume URL: <http://www.nber.org/books/fels68-1>

Publication Date: 1968

Chapter Title: Recognition Methods

Chapter Author: Rendigs Fels

Chapter URL: <http://www.nber.org/chapters/c1091>

Chapter pages in book: (p. 39 - 46)

Recognition Methods

It is hard to get evidence on the comparative success of different methods of recognizing cyclical turns, because all analysts are to some extent eclectic. Two main methods, however, can be distinguished in principle and sometimes in practice as well. Perhaps the most popular method of forecasting uses the national income framework. Separate provisional forecasts are made for the components of GNP. These are added and revised in the light of any discrepancies between the total and the assumptions underlying the provisional forecasts. Such a forecast may yield a prediction of a turning point, especially if it is a quarterly multiperiod forecast.⁴⁰ The other principal method uses the business cycle indicators. Besides these two well-recognized methods, the quantity theory of money can be and sometimes is used as a basis for forecasting turning points. And it is possible to forecast without any recognizable method at all other than interpretation of available data with the help of theory, experience, judgment, and intuition.

Since the discussions of turning points by the eight analysts in our main group were mostly informal, I cannot tell to what extent they relied on national income models, on the indicators, or on something else. For want of more knowledge and a better name, I shall refer to them as "eclectic." Two other publications, available for only part of the period under study here, relied heavily on business cycle indicators. In addition, Victor Zarnowitz has generously furnished me with data based on multiperiod forecasts of GNP (mostly quarterly). There is a presumption that Zarnowitz' forecasters relied more heavily on national income models than the two in my sample who used the indicators.

⁴⁰ The econometric method may be regarded in principle as a variant of this method, since econometric forecasts commonly use the national income framework. Because data for assessing the recognition record of the econometric method is meager, I shall not discuss it further. The accuracy of econometric forecasts is, however, being studied by Jon Cunyningham as a part of the National Bureau's project.

Data for both users of the indicators approach can be had for only four turns (1957-61). Comparison of the average scores of these two with the eight eclectics for the four turns together conceals such diversity that its value is doubtful.⁴¹ Suffice it to say that the comparison does not establish any striking differences in the success of the two groups.

Although the average scores of the two and the eight do not display marked differences for the four turns between 1957 and 1961 taken together, the same is not true for individual turns. In 1957 (see Chart I-8), the two using the indicators approach scored better on degree of certainty than the other eight in nine out of the ten months scored, and during the early months the difference was marked.⁴² The inaccurate statistical information on inventory investment in 1957 affected users of national income analysis more than those who put their faith in business cycle indicators.⁴³

Whereas the two relying heavily on the indicators scored better with respect to degree of certainty than the eight in 1957, the following year it was the other way round (see Chart I-9). In this case, a scrap of addi-

⁴¹ The two users of indicators had average scores for accuracy of dating a shade higher than the average of the eight eclectics during the three months before the four turns (between 1957 and 1961). Their forecasts of dating were markedly worse at and immediately after the turn, but during the period from three to six months after the reference dates, their scores for accuracy of dating were considerably higher. These results give no evidence that the indicators are especially helpful in determining dating contemporaneously with cyclical turns. The suggestion that the indicators have considerable value some months after the turn is subject to the qualification that many of the other analysts lacked interest in the question of dating. They were concerned with whether a peak or trough had occurred; exactly when it occurred did not matter. The very decision to use the indicators may imply a greater interest in the subject.

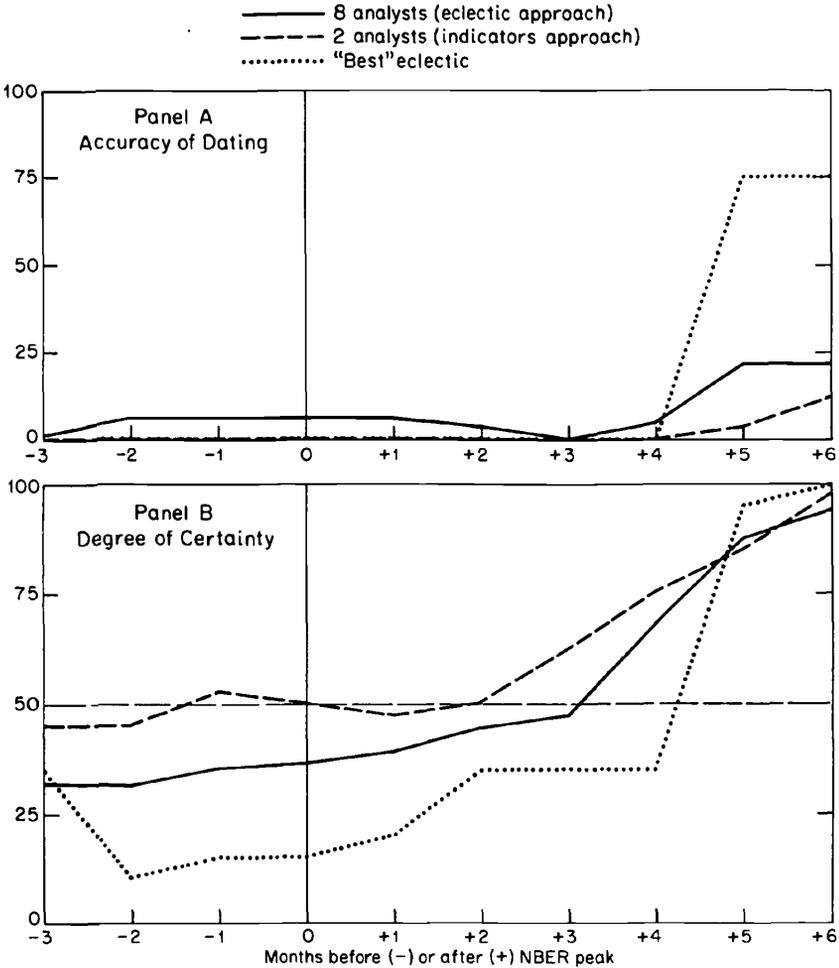
The publications relying on the indicators had average scores for degree of certainty slightly higher than the eight eclectics, but not enough higher, in view of the shortcomings of the scoring procedures, to warrant anything in the way of conclusions about recognition methods (except, perhaps, the speculation that recognition depends more on the judgment of the analyst than on the specific methods he uses).

⁴² Both groups did poorly with respect to accuracy of dating in 1957. Which did worse depends on whether the mean or the median is used for the eight, the median being zero in every single month. One of the two users of indicators had poor scores (mostly zeroes) because he consistently dated the peak much too early. The other one first predicted a date much too late and thereafter ignored the subject.

⁴³ Although inventory investment is now one of the NBER leading indicators, it was not put on the list until 1960 (Moore, *Business Cycle Indicators*, Vol. I, p. 55). What is more to the point, the indicator approach, which uses a substantial number of series without formally weighting them, is less likely to be affected by an error in any one series.

CHART I-8

Recognition Scores in the Vicinity of the 1957 Peak:
 Comparison of Indicators Approach with Eclectic Approach

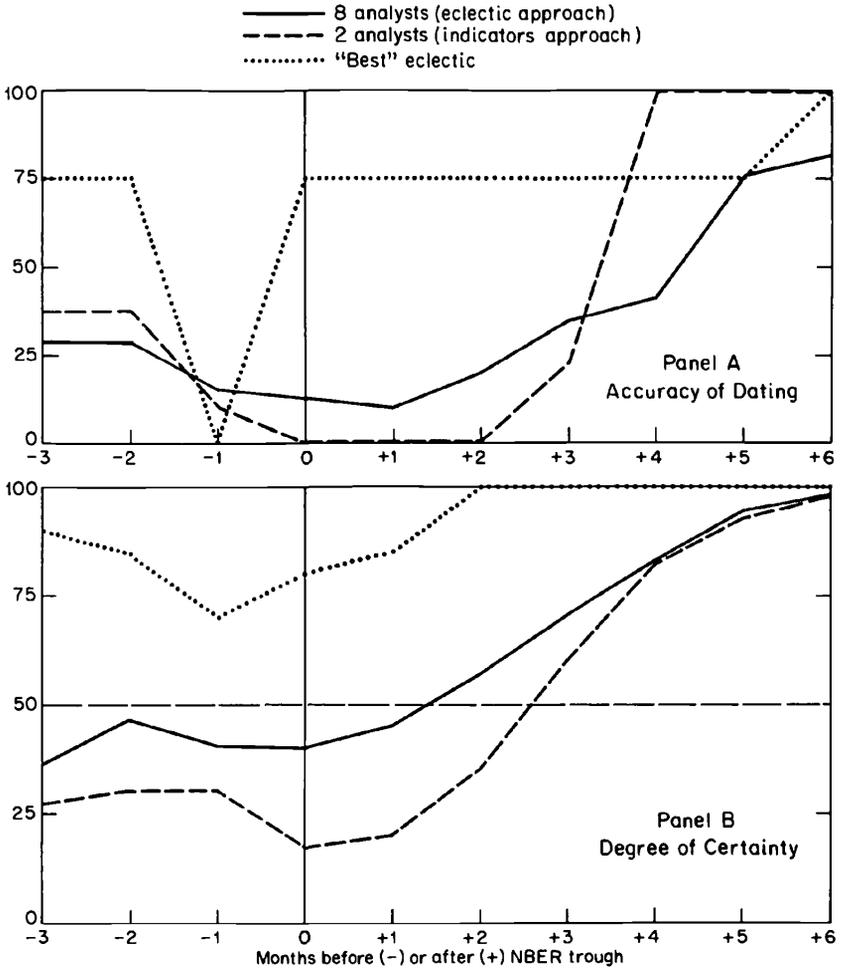


SOURCE: Appendix I, Table F.

tional information tends to support the hypothesis that a difference in method was responsible. An economist's testimony before a Congressional committee at the end of April made it clear that (1) his prediction of a weak upturn in the third quarter was based on a GNP

CHART I-9

Recognition Scores in the Vicinity of the 1958 Trough:
Comparison of Indicators Approach with Eclectic Approach



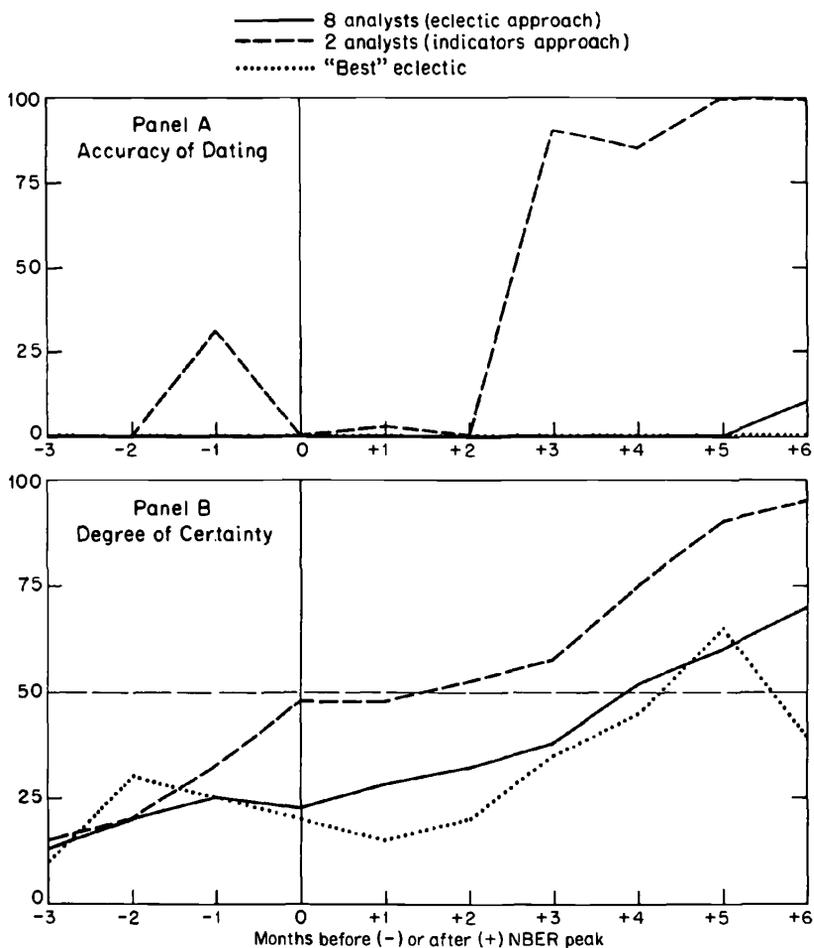
SOURCE: Appendix I, Table G.

model and (2) his low confidence that the upturn would materialize was based on the NBER indicators, which did not look good to him.

In 1960 (Chart I-10), those using the indicators again outscored the other eight, this time in accuracy of dating as well as in degree of certainty. Although in February and March they succeeded no better than

CHART I-10

*Recognition Scores in the Vicinity of the 1960 Peak:
Comparison of Indicators Approach with Eclectic Approach*

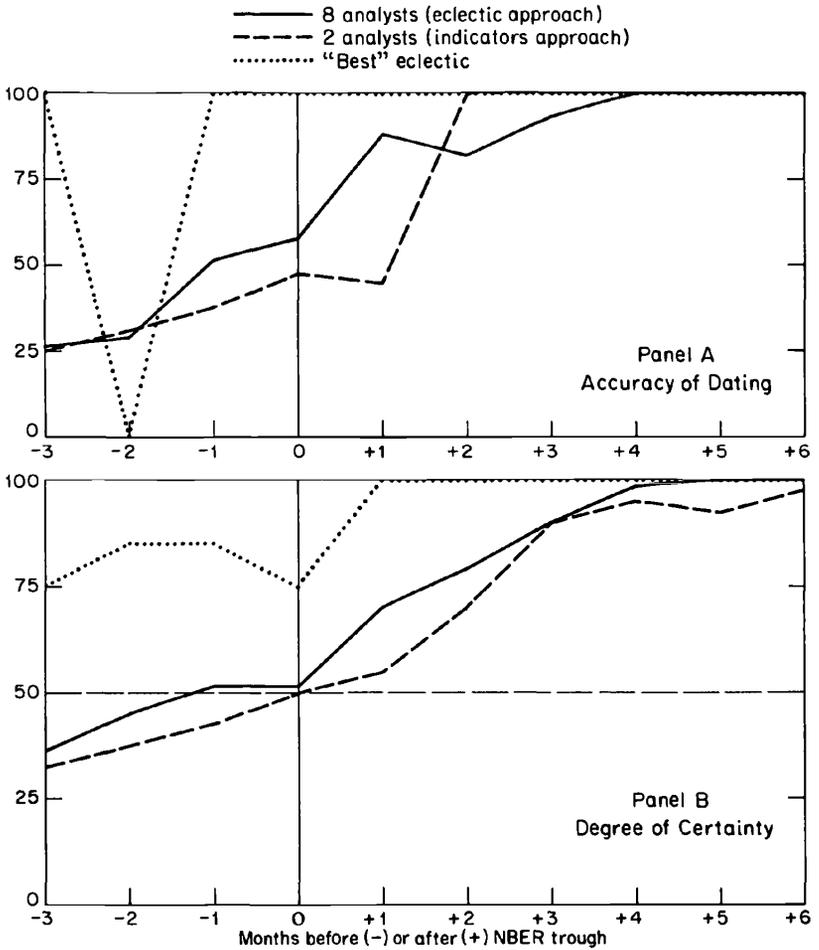


SOURCE: Appendix I, Table H.

the others in anticipating the turn, later on they were more alert to early signs that it was already past. But they did not do as well as the three other economists alluded to earlier in footnote 31, two of whom on the basis of Federal Reserve policy called the turn as early as the spring of 1959.

CHART I-11

Recognition Scores in the Vicinity of the 1961 Trough:
Comparison of Indicators Approach with Eclectic Approach



SOURCE: Appendix I, Table I.

In 1961 (Chart I-11), the eight eclectics outscored the two relying on indicators with respect to both accuracy and certainty.

In sum, if dubious evidence can be believed, the indicators seem quicker than eclectic methods to give warning of downturns and slower to herald upturns. The result may, however, reflect a difference in out-

look. The eight eclectics in general scored higher at troughs than peaks, a result consistent with the hypothesis that they inclined toward optimism. (Reading the forecasts of the "best" of the eight gives a strong impression of an optimistic bias on its part. Compare Charts I-8 and I-10 with I-9 and I-11.) Those forecasters who decided to put their faith largely in business cycle indicators may have done so because by temperament and training they were inclined to think in terms of peaks and troughs. (Reading the forecasts of one of the two gives a strong impression of a watch for the next peak as soon as the trough is past.) The two may not have shared the optimistic outlook of the others.⁴⁴ But the two may not be representative of users of the indicators, and the eight may have used indicators more heavily than appears to have been the case. Firm conclusions cannot be reached about methods where individual judgment is as important as in forecasting.

Let us turn now to a comparison of Zarnowitz' sample with mine. Zarnowitz has provided data for four sets of forecasts (A, C, D, and G, as he designates them). Two of the sets represent company forecasts of GNP. The other two result from averaging the GNP forecasts of groups of economists who made their forecasts simultaneously and on a comparable basis. Altogether there are twenty-four forecasts made in the vicinity of actual cyclical turns, yielding thirty-six observations (since these are multiperiod forecasts, the same forecast may predict—or not predict—two actual turns occurring close together). In addition, there are six forecasts yielding seven observations of predicted turns that did not occur. Since the forecasts were made at rather infrequent intervals, it is not possible to trace the pattern of increasing recognition in Zarnowitz' sample nor to make a complete comparison with mine.

I scored the thirty-six observations pertaining to actual turns for both accuracy of timing and degree of confidence. There was a minor difficulty with scoring for timing,⁴⁵ and a major one with scoring for cer-

⁴⁴ The certainty scores, which are better indicators of outlook than dating scores, bear out this speculation. The average certainty scores for the two forecasters relying on business cycle indicators were virtually the same for the two peaks together as for the two troughs, whereas for the eight eclectics the average was markedly higher at the two troughs. On accuracy of dating the two using the indicator approach scored markedly higher at the two troughs than at the two peaks, but the disparity was not as great as for the eight eclectics.

⁴⁵ For example, one forecast was for a trough in the second quarter of 1958. Since the NBER trough came in April, mechanical application of our methods would have yielded a score of 100, the highest possible. But the actual trough

tainty. All that can be inferred with respect to certainty is whether the score should be above or below 50. If there was a turn in the numerical forecasts of GNP, it could be inferred that the forecaster thought a cyclical turning point was more likely than not, and vice versa, but greater precision was not possible. Another difficulty was that a number of the forecasts in Zarnowitz' sample were made in months for which we did not score our sample.

For two or three of the observations, the average scores for my sample were higher than for Zarnowitz'.⁴⁶ In twelve cases, Zarnowitz' scored higher. In the remaining 21 or 22 cases, there was no way to tell which group deserved the higher score.

It appears, therefore, that Zarnowitz' forecasters had greater success in predicting cyclical turns than mine. Such a conclusion, however, rests on a very flimsy foundation. Besides the inadequacies of the data already mentioned, Zarnowitz has found that the mean of a group of forecasts of GNP is superior to most of the individual forecasts entering into the average. Scoring averages for his two groups (in one case the mean, in the other the median) may result in higher scores than would be obtained by averaging scores for each individual forecast in the group.

If we take at face value the conclusion that the forecasters in Zarnowitz' sample were better at recognizing cyclical turns than mine, the superiority may be associated not with a difference of methods but with the facts that (1) my group all published their forecasts, a circumstance that might induce caution about asserting a change in direction, and (2) Zarnowitz' made quantitative, multiperiod forecasts of GNP, a circumstance that kept them from taking refuge in verbal ambiguity.

Of the six forecasts of turns that did not occur in Zarnowitz' sample, one was made in December 1946 for a peak in the second half of that year, one was made in January 1956 for a peak in the first quarter and a trough in the third, and the other four were made in 1962-63. Since the indicators misled some forecasters in both 1956 and 1962-63, there is little ground here for inferring which method is superior with respect to avoiding false signals.

in GNP came not in the second quarter but in the first, so the score given was reduced to 75.

⁴⁶ In one case, a different result is obtained depending on whether my sample is deemed to consist of the eight eclectics only or of all ten publications.