Volume Title: Business Cycles: Theory, History, Indicators, and Forecasting

Volume Author/Editor: Victor Zarnowitz

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-97890-7

Volume URL: http://www.nber.org/books/zarn92-1

Conference Date: n/a

Publication Date: January 1992

Chapter Title: Composite Indexes of Leading, Coincident, and Lagging Indicators

Chapter Author: Victor Zarnowitz

Chapter URL: http://www.nber.org/chapters/c10382

Chapter pages in book: (p. 316 - 356)

11 Composite Indexes of Leading, Coincident, and Lagging Indicators

11.1 Objectives, Standards, and Assessments

11.1.1 Reasons for Combining Indicators into Indexes

In concurrence with much of the literature, this book argues that there is no *single* proven and accepted cause of *all* observed business cycles.[1] Instead, there are a number of plausible and not mutually exclusive hypotheses about what can cause downturns and contractions, upturns and expansions. Similarly, no *single* chain of symptoms exists that would invariably presage these developments. Instead, there are a number of frequently observed regularities that seem likely to persist and play important roles in business cycles but are certainly not immutable. The study of modern economic history reveals a mixture of unique and common characteristics in each recorded business cycle. It also suggests that certain systematic changes in short-term macrodynamics can be linked to long-term changes in the economy's structure and institutions and the government's role and policies.

All this combines to explain why it has proved so difficult to make progress toward a unified theory of business cycles, despite the great ingenuity of the theorists and considerable gains in tested knowledge achieved in the empirical work on the subject. Now, for the same reasons, the performance of individual indicators in any given period is apt to vary depending on which causal elements are dominant and how their working manifests itself. In particular, some leading indicators turn out to be most operative and useful in one set of conditions, and others in a different set. To increase the chances of getting true signals and reduce those of getting false ones, it is therefore advisable to rely on a reasonably diversified group of leading series with demonstrated

1. This section is based in part on Zarnowitz and Boschan 1975a, as are some other paragraphs of this chapter.

predictive potential. This suggests combining selected leaders into an appropriately constructed index and monitoring changes in that index as well as in its components on a regular basis. The argument can be readily generalized to composite indexes of roughly coincident indicators and of lagging indicators as well.

Second, the measurement errors in individual indicators are often large, especially in the most recent observations based on preliminary data. To the extent that the data errors in the different indicators are independent, the risk of being misled can be reduced by evaluating the signals, not from any one series viewed in isolation, but from a number of related series. The latter, however, must be sufficiently differentiated and not just provide different measurements for essentially the same variables. This is because multiple counting should be avoided inasmuch as it results in unintended overweighting of some elements in an index.

The third reason is an enhancement of predictive ability that can be achieved by the reduction of pure "noise." In general, indicators tend to react not only to sustained cyclical fluctuations but also to frequent disturbances of all kinds, for example, major strikes and foreign wars. This is particularly true of the sensitive leading indicators. Hence, the month-to-month changes in these series (after elimination of any seasonal elements) usually reflect the short erratic movements much more than the longer cyclical ones. By combining the series into an index, some of that noise is eliminated; that is, a well-constructed composite index can be much smoother than any of its components.

The corollary of these arguments for constructing indexes from selected series with common timing patterns is that a failure of an individual indicator does not refute the method. Rather, such a failure merely impairs and, if repeated, ultimately invalidates the particular series concerned. Unless the underlying economic process is significantly altered, the problem reduces to getting a better representation for it by improving the series or replacing it with one that would perform satisfactorily. On the other hand, even a failure on a single occasion would have strong negative implications for the indicator approach if it extended to a whole set of principal series combined in an index.

11.1.2 Criteria for the Evaluation of Indicator Performance

Historically, six criteria were applied in assessing and selecting the NBER cyclical indicators. They refer to the following questions:

1. How well understood and how important is the role in business cycles of the variables represented by the data? (The judgment on this is quantified in the score for *economic significance*.)

2. How well does the given series measure the economic variable or process in question (*statistical adequacy*)?

3. How consistently has the series led (or coincided or lagged) at business cycle peaks and troughs (*timing at recessions and revivals*)?

4. How regularly have the movements in the specific indicator reflected the expansions and contractions in the economy at large (*conformity* to historical business cycles)?

5. How promptly can a cyclical turn in the series be distinguished from directional change associated with shorter, irregular movements (*smoothness*, which is inversely related to the degree of statistical noise)?

6. How promptly available are the statistics and how frequently are they reported (*currency* or timeliness)?

A formal, detailed weighting scheme according to the above criteria was first developed and applied in Moore and Shiskin 1967. In the Zarnowitz and Boschan studies (1975a, 1975b, 1975c), a revised version of the same approach was used. The scores for each of the six major characteristics and their many relevant components were computed on the 0-to-100 scale and then combined into a total score by means of the weights.[2]

The evaluation of economic significance is difficult and inevitably subjective; hence much of it was handled by preselection, with the minimum acceptable score set at 70%. The scores for all other criteria were based on essentially objective statistical measures. Thus, the quality of the reporting system is assessed according to whether it is set up directly for statistical purposes, is a by-product of an administrative system, or is nonexistent (as for series estimated indirectly from related variables). Other aspects of statistical adequacy include the coverage of process (full enumeration, probability sample, other) and of time period (full month, or 1 week, or 1 day per month, etc.); the availability of estimates of sampling and reporting errors; the length of the series and comparability over time (breaks are penalized); and the frequency of revisions (none, once a reporting period, or more often). More recently, the revisions were given a more elaborate treatment and a separate score reflecting their relative size as well as frequency.[3]

The cyclical timing performance of an indicator is appraised mainly according to the probability that the observed number of timing comparisons of a given type will be equaled or exceeded by chance. The dispersion of the leads and lags about their means is also taken into account but with a smaller weight (of about 20% of the total timing score). The leads and lags of the series are

2. In 1967, the first four criteria received weights of 20% each; the last two received weights of 10% each. The results referred to the period before 1966 (as far back as the data were available but with a heavy preponderance of the evidence from the post–World War II years). The timing of the indicators at peaks and troughs was handled symmetrically.

In 1975, criteria 1, 2, and 4 received weights of $1/6$ each, timing $4/15$, smoothness $2/15$, and currency $1/10$. The results referred to the period 1947–70. An asymmetrical distribution of the cyclical timing comparisons at peaks and troughs of the postwar period was applied. Cf. Moore and Shiskin 1967, pp. 22–27, and Zarnowitz and Boschan 1975a, pp. 2–4.

3. For more detailed information on scoring for statistical adequacy, errors and revisions, smoothness and currency, see chapter 13 below. Also, on these and the other scores for the components of the 1966 and 1975 indexes, see Moore and Shiskin 1967 and Zarnowitz and Boschan 1975a and 1975b.

measured at the NBER-established references dates for business cycle peaks and troughs separately.

Conformity is evaluated by relating the number of business cycle phases that are matched by the specific-cycle movements (expansions and contractions) in the given series to the total number of phases covered and then computing the probabilities for the observed records. Other elements of the conformity score are the frequencies of "extra" movements in the indicators that do not match the phases of general business fluctuations and can result in misleading "false signals." The amplitudes of cyclical changes in the series are also accounted for here (the larger and more distinct movements score higher).[4]

Smoothness depends on the relationship between the irregular and the trend-cycle components of the series adjusted for seasonal variation. Large erratic variations are common among many leading indicators and constitute the main defect in some. Smoothing the data with short, trailing moving averages can be advantageous for some erratic series that have long leads and high currency scores (i.e., are compiled frequently and released promptly). To put it differently, trade-off relationships exist between the smoothness, currency, and timing characteristics (e.g., moving averages increase smoothness but reduce currency and possibly the lead times).

11.1.3 The Indicator Scores and Selection of Index Components

Table 11.1 presents means and standard deviations of each of the principal component scores and of the total scores for 108 cyclical indicators that were regularly presented in *BCD*. These statistics refer to data for 1947–80. The series are classified by cyclical timing (pt. A) and by economic process (pt. B). Listed in part A are also the scores of the composite indexes of leading, coincident, and lagging indicators as well as of the groups of series included in each of these indexes.

The major purpose of the original scoring efforts in the mid-1960s and again in the early 1970s was to help select the most consistently cyclical series for the indicators section of *BCD*. This explains in large part why the overall average scores in the table (col. 8) are all relatively high and clustered. They do not differ greatly across either the economic process or the cyclical timing groups.

It is also to be noted that the differences between the component scores are generally large and often significant but that they offset each other to a large extent. For example, the leaders as a group rank well below the coinciders with respect to smoothness, but the reverse obtains for revisions (see lines 1–2 and 5–6). Production and income series (lines 20–21) score high on timing

4. Within the total score for conformity, probability gets a maximum of 50, extra turns 30, and amplitude 20 points.

Table 11.1 Means and Standard Deviations of Scores of Cyclical Indicators, 1947–80

Line	Timing (1)	Conformity (2)	Smoothness (3)	Currency (4)	Statistical Adequacy (5)	Revisions (6)	Economic Significance (7)	Total (8)
			A. CLASSIFIED BY CYCLICAL TIMING (108)					
			All Leading Series (47)					
1	76	68	57	57	78	72	74	71
2	11	18	27	29	12	22	7	6
			Components of the Leading Index (12)					
3	79	72	68	80	75	65	74	74
4	10	16	20	10	8	30	9	5
			All Roughly Coincident Series (18)					
5	86	77	87	57	80	52	82	78
6	18	7	15	28	8	18	11	8
			Components of the Coincident Index (4)					
7	95	82	95	74	77	40	88	82
8	5	4	10	13	5	16	5	6
			All Lagging Series (26)					
9	81	72	85	62	75	82	78	77
10	16	15	17	32	13	20	5	6
			Components of the Lagging Index (6)					
11	86	71	87	75	73	70	80	78
12	6	18	16	18	15	24	6	8
			Unclassified (17)					
13	75	77	78	72	74	75	76	75
14	18	16	22	26	14	18	8	8
			Composite Indexes (3)					
15L	86	72	100	80	77	0	90	75
16C	100	86	100	80	77	20	90	83
17Lg	92	86	100	80	73	0	90	78
			B. CLASSIFIED BY ECONOMIC PROCESS (108)					
			I. Employment and Unemployment (15)					
18	84	85	81	80	71	83	81	81
19	9	9	16	0	13	18	10	5
			II. Production and Income (10)					
20	93	80	92	56	74	46	84	79
21	9	6	10	31	11	16	8	7
			III. Consumption, Trade, Orders, and Deliveries (13)					
22	75	72	71	71	81	63	70	73
23	22	15	21	18	7	18	10	9
			IV. Fixed Capital Investment (18)					
24	80	73	68	52	79	67	74	73
25	15	12	23	30	9	22	5	4
			V. Inventories and Inventory Investment (9)					
26	77	64	71	50	77	64	77	70
27	15	17	38	11	9	17	5	6

Table 11.1 **Continued**

Line	Timing (1)	Conformity (2)	Smoothness (3)	Currency (4)	Statistical Adequacy (5)	Revisions (6)	Economic Significance (7)	Total (8)
			VI. Prices, Costs, and Profits (17)					
28	76	61	72	40	73	75	75	69
29	17	15	14	32	9	22	5	6
			VII. Money and Credit (26)					
30	77	68	65	70	82	82	78	75
31	14	19	33	30	17	19	5	7

Source: U.S. Department of Commerce, Bureau of Economic Analysis 1984, table 7, pp. 169–71.
Note: In each section, except Composite Indexes (lines 15–17), entries in the first line are means of the scores of the individual series, and entries in the second line are the corresponding standard deviations. Entries in lines 15–17 are the mean scores of the leading (L), coincident (C), and lagging (Lg) composite indexes. Series are seasonally adjusted except for those that appear to contain no seasonal movement. The number of series in each group is shown in parentheses.

(which is here typically coincident) but low on currency and revisions (the data from GNP accounts are quarterly and subject to several and often large alterations). The series relating to consumption and trade do worse on timing, conformity, and smoothness but better on currency, statistical adequacy, and revisions (lines 22–23). In sum, the dispersion across the different groups of indicators is much greater for the average scores of the individual attributes than for the average total scores. The standard deviations within the groups, too, tend to be much larger for the component than for the overall scores.

The indicators included in the leading, coincident, and lagging indexes were chosen from the best-scoring series in the respective timing categories. Obviously, high consistency of procyclical or countercyclical behavior, that is, high conformity or coherence (correlation with business cycles, allowing for any systematic leads or lags), is a desirable characteristic of any index component. Timing consistency is particularly important. In addition, prompt availability of reasonably accurate data is an essential requirement for an index that is to be used in current business analysis and forecasting. This means that the component series should be released frequently with short publication lags. Thus, beginning with the 1975 list, the composite indexes of the U.S. Department of Commerce include monthly series only.[5]

As implied by these observations, the components of the leading index per-

5. Dropping GNP from the coincident index and corporate profits from the leading index involved very difficult decisions in view of the importance of the represented variables. Real GNP is, of course, the central measure of total output, but the data are quarterly and much revised. Important studies of business cycles (notably Mitchell 1913, 1927) ascribe a major role to profits; also, there is evidence of a strong tendency for total corporate profits to lead. But few concepts are more difficult to measure than profits in the true economic sense. Data on corporate profits are compiled only on a quarterly basis and are available only with long delays and sequences of revisions.

form better than the average leading series on timing, conformity, and, particularly, smoothness and currency. However, this is partly paid for by lower scores on revisions and (to a lesser extent) statistical adequacy. Due to these offsets, the overall advantage of the index component series is modest (cf. lines 1 and 3). Similar statements can be made about the coincident and lagging series (cf. lines 5 and 7, and 9 and 11). Note that larger differences would be shown had the index components been excluded from the "all series" groups.

The indexes themselves score as well or better than the averages of the corresponding component series in all categories except revisions. The gains from combining the indicators are considerable for timing, conformity, and currency but much larger yet for smoothness, especially in the case of the leading indicators. However, until very recently, the indexes earned zero or very low scores for revisions, mainly because each of them contained some series that were not available in time to be included in the first release. No estimates of the contributions of the missing components were made, so that a month later, when the lagging data first appeared, the indexes typically showed large revisions. Hence, the advantages elsewhere were again largely dissipated here (cf. lines 3 and 15, 7 and 16, and 11 and 17). The two tardy components were eliminated from the Department of Commerce leading index in 1989 with the intention of radically reducing the size of the revisions in that index (Hertzberg and Beckman 1989; for detail, see table 11.3 and text below).

Clearly, more is required of a good index than of a good single indicator. For reasons already stated, the indexes should be reasonably diversified in their economic coverage in order to have the potential to function well in the longer run or under widely varying conditions. Also, each different variable of interest should be given the best and broadest representation possible. Consequently, the selection of index components favored comprehensive series drawn from all major economic process groups that fit into the given timing pattern. At times, the implementation of this principle involved some costs in terms of lower scores than would be available otherwise (i.e., from collections of series less diversified and/or less aggregative).

The existing indexes are designed to lead (or coincide or lag) at both peaks and troughs of business cycles. This is restrictive since it excludes series with mixed timing patterns, which prevail in some areas for reasons that are well understood (see chapter 10, sec. 6). Such series (i.e., those that lead at peaks and lag at troughs or vice versa) are included in the "unclassified" group in table 11.1 (lines 13–14). Note that these indicators score on average about as well as the other groups.

Recall also that in recent business cycles leads have been much more common, but also much more variable, at peaks than at troughs. As shown in table 11.2, the timing scores averaged considerably less at peaks than at troughs for most indicators. Of the timing groups, only the laggers score slightly higher

Table 11.2 Mean Scores of Groups of Cyclical Indicators and Composite Indexes, Timing at Peaks and Troughs, 1947–80

Classified by Timing[a] (no. of series) (1)	Mean Timing Score: Peaks (2)	Mean Timing Score: Troughs (3)	Classified by Economic Process[b] (no. of series) (4)	Mean Timing Score: Peaks (5)	Mean Timing Score: Troughs (6)
All L series (47)	48	78	I (15)	58	85
L I components (12)	62	79	II (10)	73	87
Leading index (1)	76	94	III (13)	41	76
All C series (18)	57	81	IV (18)	49	82
C I components (4)	53	91	V (9)	57	69
Coincident index (1)	96	97	VI (17)	44	74
All Lg series (26)	76	71	VII (26)	66	65
Lg I components (6)	88	83			
Lagging index (1)	94	82	All series (108)	55	76
Unclassified (17)	46	76	All index components	67	82

Source: See table 11.1.

[a]Abbreviations: L = leading; I = index; C = roughly coincident; Lg = lagging.

[b]Groups are identified by roman numerals as in table 11.1. See table 11.1 for the titles of these economic process groups.

at peaks than at troughs, and of the economic process groups only the money and credit series (VII) do so. However, the timing performance is much better and more balanced between peaks and troughs for the composite indexes than for the corresponding "all series" groups.

11.1.4 Measurement Errors

Macroeconomic indicators are obtained by processing and aggregating primary data that as a rule contain a variety of errors, both random and systematic.[6] The errors in the components may either offset or reinforce each other in the aggregate. The procedures for the derivation of economic macromeasures can themselves contribute to the inaccuracy of the resulting time series. The more complex the process and the more numerous the estimations and approximations that are involved, the greater are the chances of substantial defects, with conceptual, procedural, and statistical ingredients that are difficult to isolate and assess.

As shown by the pervasiveness and persistence of revisions in economic statistics, some of the errors are of the kind that can be reduced only partially

6. Systematic errors may arise from nonprobability sampling or inadequate sampling or faulty enumeration. Respondents and/or collectors of the data may be poorly informed or poorly motivated or both. Concealment and falsification are possible. In short, the data vary greatly in coverage and quality, depending on the knowledge and cooperation of the people who make them. For further discussion of these and other matters covered in this section of the text and the next one, see Zarnowitz 1982c.

and gradually. To be prompt and frequent, as demanded for current business analysis and forecasting, information on important indicators must often take the form of preliminary figures subject to repeated and possibly large alterations. The discrepancies between the successive releases of a time series represent for the most part errors resulting from large lags in the availability of primary data. Although other errors created by conceptual and procedural problems may well be more serious, they tend to be less identifiable. Their detection and correction occurs only sporadically, as a result of benchmark revisions and definitional changes that cause "breaks," that is, discontinuities in several of the major indicator series.

Some of the cyclical indicators included in the BEA indexes have only occasional and mostly minor revisions (average workweek, layoff rate, unemployment duration) or no revisions at all (stock prices, vendor performance, the prime rate). However, for most of the index components the first preliminary figure A_1 (issued in the current month for the previous one) is revised 1–4 times in as many successive months (A_i, $i = 2, 3, 4$). The first two changes generally account for a very high proportion of the total revision.[7]

Let the error term in successive revisions be defined as $E_{it} = A_{it} - A_{1t}$. For most indicators, both the averages taken without regard to sign and the standard deviations of E_{it} tend to increase with i, the index of the time distance between the revisions. The mean errors are predominantly negative, suggesting underestimation of levels in the early data, but they are mostly small and independent of i. Tests that $a_i = 0$ and $b_i = 1$ in the regression equations $A_{it} = a_i + b_i A_{1t} + e_{it}$ produce mixed results. Although for most indicators the F- and t-ratios are not clearly significant, they are so for some series that are based on indirect estimation and, in their early version, seriously incomplete data. The measurement errors revealed by the revisions are not systematically larger in the leading than in the coincident or lagging indicators. The revisions are typically fewer and less diffused over time in the monthly indicators than in quarterly aggregates from the national income and product accounts.[8] However, very large and frequent revisions are found in both sets of time series for such important variables as inflation, monetary changes, inventory investment, and profits.[9] Adjustments for price changes and other procedures that are inevitably intricate and approximate in nature tend to add to the probable error.

For the composite indexes of cyclical indicators, incomplete coverage in

7. For quantitative evidence, see Zarnowitz 1982c, pp. 93–103. This and the following paragraph present in brief the main qualitative findings from this study.

8. The GNP revisions extend from those in each of the three months following the quarter to which the data refer, through those in July of each of the three successive years, to the large but infrequent benchmark overhauls that are important in the historical sense rather than for any practical current purposes. See Zellner 1958; R. Cole 1969b.

9. In recent times, series for various monetary aggregates have been published at monthly and weekly intervals and revised almost as frequently (see, e.g., Pierce et al. 1981). On the importance of revisions in the GNP implicit price deflator, see Keane and Runkle 1990, pp. 722–24. On measures of inventory investment and corporate profits, see Zarnowitz 1982c, pp. 95–96.

the early releases has long presented a major practical problem. In the absence of good and timely estimates for the missing components, this problem is a particularly serious one for series to be used in forecasting such as the leading indexes. Further difficulties in the application and testing of these series arise from occasional discontinuities due to changes in the composition, weights, or other technical aspects of the indexes. (About all of this more will be said later in this chapter.)

Measurement errors are very common in economic data and they affect economic behavior, analysis, and forecasting; yet their role is not well understood. For example, consider tests of the predictive value of the leading index (l) based on forecasts of real GNP (q) obtained from lagged values of q and l. Should l be represented by preliminary values because these alone are available to an authentic ex ante forecaster? And should q be represented by final values because these alone represent the "truth" to be predicted? To answer yes to both questions implies that the forecasts are expected to eliminate correctly the cumulative future errors in the series on real GNP growth (as revealed by the whole string of statistical and conceptual revisions). But these errors may be both highly significant and to a large extent unforecastable. Moreover, the lagged values of q in their final form are unknown to the real-time forecaster. In sum, tests of this type could be severely biased against finding l a good predictor of q because the measures used leave large errors in l but not in q.[10]

Testing forecasts against revised instead of initial values of the target series *may* result in a spurious rejection of the hypothesis that the forecasts are rational or unbiased (cf. Keane and Runkle 1990). However, the errors in the predicted variable can have different consequences, depending entirely on their nature. If the errors are small and random, they will matter little or not at all. If they are significant and systematic but largely predictable, they should optimally be taken into account by the forecaster (see chapter 13).

On the assumption that the revisions, as intended, cumulatively improve the data,[11] they provide important information: the larger, the less stable, and the more stretched out in time they are, the less dependable are the most recent statistics relating to the current economic situation. For series that are not revised, this information is simply lacking: they may or may not have significant measurement errors. The absence of revisions, or the cessation of further revisions, does not prove the absence of errors.

Measures of statistical adequacy help assess the quality of a time series but offer no substitute for the quantitative estimates of error available from the revisions. For 110 *BCD* indicators as of 1975–80, scores for these character-

10. Diebold and Rudebusch 1990b follow a method analogous to that described in the text paragraph above, but their target variable is industrial production, which is not as strongly affected by errors that can be inferred from data revisions as real GNP is. See section 11.3.4 below.

11. This is not always true: it is not unusual to find some revisions, in a chain of several, which increase, rather than reduce, the discrepancy between the prior and the final estimate.

istics averaged 70 (out of 100), fell heavily in the 60–79 range, and were not systematically differentiated by broad categories of economic process and cyclical timing.[12]

11.1.5 Information Lags

Fourteen of our cyclical indicators are weekly or daily series that can be smoothed with a minimum loss of timeliness and provide monthly estimates the same month or early in the next month; 67 are monthly series available with lags of 1 month (48), 2 months (15), and 3 month (4); and 29 are quarterly series available some time during the next quarter. The currency scores are 100 for a series that is collected at least weekly, 20 for a quarterly series, and intermediate for monthly series—the lower the score, the longer the lag of release.

Even if a series is available promptly, its month-to-month movement may be so obscured by either seasonal change or irregular variation (noise) as to shed little light on the longer, cyclical movements and trends that are of primary interest for current business analysis and forecasting. To the extent that the seasonal fluctuations are distinct, independent, and stable, they can be reasonably well measured and eliminated (most indicators are presently reported in seasonally adjusted form). The noise element varies greatly across the indicators, as shown by measures computed for each of them in the Bureau of the Census X-11 program of decomposition and seasonal adjustment. The ratio $\bar{I}_i/\bar{C}_i$ compares the irregular to the trend-cycle component of the given series, the average changes in the two being measured over i unit periods. As the span i is increased, $\bar{C}_i$ builds up while $\bar{I}_i$ shows little (and no systematic) change, so the ratio generally declines. The shortest span i, in months, for which $\bar{I}/\bar{C} < 1$ is called MCD (months for cyclical dominance). The smoother a series, the smaller are its $\bar{I}/\bar{C}$ and MCD values. In a set of 81 weekly and monthly indicators, 30 series had MCD = 1, 21 series had MCD = 2, and 30 series had MCD$\geq$3 (37%, 26%, and 37%, respectively). For 29 quarterly indicators, the corresponding frequencies were 4 (14%), 13 (45%), and 12 (41%).

A cross-classification of the series by scores for currency and smoothness shows no tendency for the two ordinal scales to either agree or disagree with each other systematically (see Zarnowitz 1982c, table 4 and pp. 108–10). The coefficient of rank association is here $G = .07$, and the probability of the true value of G being different from 0 seems extremely low (Goodman and Kruskal 1954, 1963). Across the series, then, our measures of the data-release lag (currency) and the signal-detection lag (smoothness) are uncorrelated. By adding these lags for each series, one can obtain a rough estimate of the total lag involved in extracting information from the data.

12. See Zarnowitz 1982c, text and table 3, pp. 103 and 106–8. The series show more dispersion within than across such categories.

Table 11.3 Estimated Information Lags for 110 Cyclical Indicator Series

Information Lag (months) (1)	Weekly or Monthly (no.) (2)	Quarterly (no.) (3)	All (no.) (4)	All (%) (5)
2	22	0	22	20.0
3	25	0	25	22.7
4	18	4	22	20.0
5	6	13	19	17.3
6	3	11	14	12.7
7	4	1	5	4.6
8	1	0	1	0.9
9	2	0	2	1.8
Total	81	29	110	100.0
Average lag (months)	3.6	5.3	4.1	. . .

Source: Zarnowitz 1982c, text and table 4.

Note: The information lag is the sum of the data-release lag (1–3 months) and the signal-detection lag (1–6 months). The information lags of 2–4 months consist of data-release lags of 1–2 months and signal-detection lags of 1–3 months; the longer information lags include data-release lags of 3 months and signal-detection lags of 4–6 months.

Table 11.3 shows that the lags so calculated vary from 2 to 9 months for the weekly and monthly indicators and from 4 to 7 months for the quarterly indicators. The 41 series (37% of the sample) that have lags of 5 or more months account for about half of the sensitive indicators of fixed and inventory investment, prices, costs, profits, money, and credit. The estimates are conservative in the sense that they ignore whatever could be done to shorten the information lags by skillful data analysis (e.g., use of weekly and monthly data to correct deficit or anticipate tardy monthly and quarterly information) and forecasting (e.g., projections of lead-lag relationships). On the other hand, they are understated because the signal-detection lags are derived from revised observations rather than current figures, which are frequently preliminary and contain more statistical noise. On average, the underestimation is probably 1 or 1.5 months for all the series covered but higher, perhaps 1–3 months, for those series with significant revisions.[13]

Incomplete or defective information can contribute to expectation and decision errors. Furthermore, the longer and more staggered the information lags, the longer will such errors persist and the costlier will be their effects, detection, and correction. The actual lags and errors are unlikely to be a major source of business cycles but can surely play a significant role in their propagation (see chapter 2, secs. 2.4.3 and 2.4.5).

13. These are rough estimates, based on simple methods applicable to large volumes of data at relatively low costs, but they yield generally plausible results. They are, to my knowledge, not contradicted by other studies that use different techniques posing higher data requirements (for further detail and references, see Zarnowitz 1982c, pp. 111–12).

Informational lags and errors impede learning and prediction, yet to counteract them, longer and better forecasts are required. An effective reduction of the time needed to assess the state of the economy in the present and near future is itself a modest but significant forecasting achievement. The leading and confirming indexes can help the forecaster advance further along this way.

11.2 The Composition, Timing Records, and Construction of Composite Indexes

11.2.1 The Present Indexes

The currently published indexes of leading, coincident, and lagging indicators incorporate revisions that the Commerce Department (BEA) introduced in February 1989. Table 11.4 presents a summary of the timing records of these indexes and their components. The measures refer to the seven peaks and seven troughs of business cycles in 1953–82.

The leading index consists now of 11 series drawn from five economic process groups (it includes no series on production or income and none related to inventories). Three cyclically sensitive and important variables that were represented in the past are no longer: corporate profit, the flow of credit, and inventory investment. This is regrettable and presumably due to the deficiencies of the data, which one greatly wishes were remedied. The inclusion of quarterly or tardy monthly series would necessitate either releasing the index with long delays, using some monthly proxies first, or adding the data for these variables only after the initial estimate of the index had been published (at the cost of large index revisions). None of the alternatives are appealing. As a result, the index is somewhat less broadly based than some of its predecessors (see table 11.5 and text below). On the other hand, the new index has the advantage that all of its components are promptly available and subject to smaller revisions than the excluded series. This should make the successive estimates of the index for any month less variable and therefore easier to interpret with some confidence.

Two of the leading series on the 1989 list (*BCD* code numbers 1 and 5) are early indicators of labor market developments; three (8, 20, 92) are built wholly or partly from data on new and unfilled orders received by manufacturers; two (19, 99) are indexes of cyclically sensitive prices of assets (common stocks) and inputs (industrial materials); and 1 each represent residential investment commitments (29), the speed of deliveries (32), consumer expectations (83), and the deflated money stock (106). Except for the newly added index of consumer expectations, all these indicators stand for variables that in one form or another were given some attention in earlier vintages of the leading index as well. The coverage still relies heavily on manufacturing, construction, and selected market and trade data. Efforts to find and add leading indicators of services and international activity have so far met with little success (the change in the consumer price index for services is included in the new composite index of lagging indicators). The available time series in these areas are generally dominated by longer trends and not very cyclical.

Only the future can reveal the predictive potential and usefulness in application of the new indexes and their components, but it is somewhat reassuring to know that at least in the past the cyclical timing of the selected series has been fairly consistent. For the 11 leading indicators, the mean (median) leads at peaks had a range of 7–16 (8–15) months. The standard deviations varied from about 3 to 13 and were smaller than the corresponding means in each case except one (index of consumer expectations). At troughs, the leads were much shorter, averaging 1 month or less for four of the series and 2–5 months for the others. Here the dispersion (S.D.) measures show the same order of magnitude as the means. In other words, the leaders at peaks often had only very short leads or a more nearly coincident timing at troughs (compare cols. 1–3 and 4–6 in the first section of table 11.4). The means (S.D.) of the leads of the composite index itself were approximately 10 (6) and 5 (4) months at business cycle peaks and troughs, respectively (the median lead at troughs was only 2 months).

It will be recalled that the peak-trough differences in timing extend to a great many cyclical indicators (chapter 10). Whereas the leads tend to be much longer at peaks than at troughs of the postwar business cycles, the lags tend to be longer at troughs than at peaks. Indeed, the latter shows up as a clear contrast for the present lagging index of the Commerce Department (BEA) and for each of its seven components (see the last section in table 11.4). In short, the observations for the principal indicators and indexes reflect a more broadly observed phenomenon.

The timing differences are very likely related to the asymmetries of postwar business cycles. We are already familiar with the relevant facts: expansions have typically been much longer and more gradual than contractions. Long-term growth accounts for a part but not all of the apparently strong asymmetries (see chapter 8, sec. 8.3.4). The results of the indicator analysis suggest that separate leading indexes for peaks and troughs could have some significant advantages. In 1975, experimental indexes of this type were constructed with partially positive results (see Zarnowitz and Boschan 1975a, n. 23). They met with the criticism of being too complicated and costly to maintain and use, but recently a very different analysis gave independent support to the same idea (Diebold and Rudebusch 1989).[14]

None of the three composite indexes failed to match any of the recorded 14 business cycle turns of 1953–82, and the same is true of 17 of their component series (table 11.4, col. 7). The series real M2 missed 6 turns, 1 lagging series missed 4, and 3 series (one in each timing category) missed 2 turns, or

14. "[T]he use of two indexes, an 'expansion index' and a 'contraction index,' constructed with different components and component weights, could enhance predictive performance" (Diebold and Rudebusch 1989, pp. 386–87).

Table 11.4 **Leading, Coincident, and Lagging Indexes and Their Components, Summary Measures of Timing at Business Cycle Peaks and Troughs, 1953–82**

	Leads (−) or Lags (+), in Months							
	B.C. Peaks			B.C. Troughs			B.C. Turns Missed	Extra S.C. Turns
	Median (1)	Mean (2)	S.D. (3)	Median (4)	Mean (5)	S.D. (6)	(7)	(8)
Leading indicators								
1. Average weekly hours, manufacturing	−10	−9.7	6.6	−1	−1.0	1.0	0	4
5. Average weekly initial claims (inverted)	−11	−10.1	7.0	0	−0.1	2.0	0	4
8. Manufacturers' new orders, consumer goods and materials[a]	−13	−12.1	6.8	−1	−1.4	2.5	0	4
32. Vendor performance, slow deliveries	−9	−10.0	9.4	−4	−4.4	4.2	0	4
20. Contracts & orders, plant & equipment[a]	−8	−7.0	4.2	−1	0.1	4.1	0	2
29. Building permits, new private housing units	−11	−15.0	7.8	−3	−5.4	4.9	0	4
92. Change in manufacturers' unfilled orders[a]	−12	−12.9	7.4	−2	−3.0	3.2	0	6
99. Change in sensitive materials prices[b]	−8	−9.3	8.9	−2	−2.7	2.3	0	6
19. Index of stock prices, 500 common stocks	−9	−9.7	2.6	−4	−4.7	1.8	2	10
106. Money supply, M2[a]	−15	−16.0	6.7	−2.5	−3.5	2.4	6	2
83. Index of consumer expectations	−9	−11.6	12.6	−4	−3.9	3.1	0	4
910. Composite index of 11 leading indicators	−8	−9.7	6.1	−2	−4.6	4.1	0	4
940. Ratio, coincident index to lagging index	−12	−14.6	6.8	−1	−2.6	3.7	0	4

Coincident indicators								
41. Employees on nonagricultural payrolls	+2	+1.4	4.9	+1	+0.9	1.1	0	0
51. Personal income less transfer payments	+0.5	+1.0	1.3	−1	−1.0	0.9	2	0
47. Index of industrial production	0	−1.3	2.6	0	0	0.6	0	2
57. Manufacturing & trade sales[a]	−3	−4.3	3.3	−1	−1.1	1.8	0	0
920. Composite index of 4 roughly coincident indicators	0	−1.6	2.3	−0	−0.6	1.1	0	0
Lagging indicators								
91. Average duration of unemployment (inverted)	+1	0	3.6	+8	+9.4	4.9	0	0
77. Ratio of manufacturing and trade inventories to sales	+9	+9.9	4.4	+14	+18.4	13.0	0	4
62. Change in index of unit labor cost, manufacturing	+6	+7.3	4.7	+9	+9.4	2.2	0	10
109. Average prime rate charged by banks	+3	+4.3	3.2	+14	+17.9	19.0	0	4
101. Commercial and industrial loans[a]	+5	+5.7	5.9	+9.5	+9.8	6.0	2	2
95. Ratio of consumer installment credit to personal income	+5	+5.6	3.1	+7	+6.8	3.8	4	4
120. Change in consumer price index for services[b]	+3	+1.8	6.4	+5	+8.3	9.4	0	6
930. Composite index of 7 lagging indicators	+3	+4.7	3.8	+7	+8.4	4.9	0	2

Source: U.S. Department of Commerce, Bureau of Economic Analysis, *Business Conditions Digest* (*BCD*), December 1989, p. 104.

Note: The numbers preceding the titles of the series are the *BCD* series numbers. The list of indicators used in this table is the one introduced by the Bureau of Economic Analysis in February 1989 and used presently (see Hertzberg and Beckman 1989). The table is based on leads or lags of specific peak or trough dates that mark the cyclical turning points in the individual indicator series at the corresponding reference peak or trough dates that mark the cyclical turning points in the overall business activity. Seven business cycles are covered, with the following NBER reference dates: peaks, 7/53, 8/57, 4/60, 12/69, 11/73, 1/80, and 7/81; troughs, 5/54, 4/58, 2/61, 11/70, 3/75, 7/80, 11/82.

Abbreviations: B.C. = business cycle; S.C. = specific cycle; S.D. = standard deviation.

[a]1982 dollars.

[b]Smoothed by an autoregressive moving-average filter developed by Statistics Canada.

1 phase, each. Earlier indexes had similarly good records in the sense of having missed very few business cycle turns. Of course, it helps that this is a retrospective analysis using currently available, revised data; spotting the turns ex ante in preliminary data is at times much more difficult. Even so, missing the turning points of business cycles is definitely not the most important type of error for these indicators; giving false signals is.

This is readily seen from the contrast between the rare instances of business cycle turns missed and the high frequencies of extra specific-cycle turns missed (cf. cols. 7 and 8). Each of the leading series had some cyclical movements of its own that were not correlated with the general economic fluctuations. The modal number of extra turns was 4, some series had 2 and 6, but one (the stock price index) had 10. Interestingly, a very similar statement can be made about the lagging indicators, except that one of these had no extra turns (one, the index of unit labor cost, had 10).

As would be expected, the four comprehensive coincident indicators show a record of almost one-to-one correspondence with business cycle peaks and troughs, that is, a minimal number of errors of either type. About one third of the individual comparisons at peaks and half of those at troughs were exact coincidences (0), with the rest divided between mostly short leads and lags. Employment was slightly lagging.

The coincident index itself led at three peaks, lagged at two troughs, and had the timing 0 at each of the other nine business cycle turns. The ratio of this index to the lagging index (which reflects inversely the role of costs of labor, inventory, and finance; see chapter 10) tended to lead, by long intervals at peaks and short ones at troughs.

11.2.2 Five Successive Lists of Leading Indicators, 1950–89

The first list of NBER business cycle indicators was based on a study of nearly 500 series that varied in length but ended in 1933 (Mitchell and Burns 1938). It included 71 series "tolerably consistent in their timing in relation to business cycle revivals" and a subset of 21 "most trustworthy" of these indicators. The next review used about 800 series through 1938, classified them by timing, and selected indicators of recession as well as revival (Moore 1950). After another decade, it was possible to analyze still more series and extend the measures through 1958 (Moore 1961). Revised short and supplementary lists of indicators were shown, along with the first composite index of leading indicators. In 1966 another study introduced an explicit scoring plan and used it to evaluate more than a hundred series (Moore and Shiskin 1967). The results included new long and short lists of U.S. indicators classified by cyclical timing and economic process, and corresponding composite indexes.

In 1972–75, the BEA conducted a comprehensive appraisal of cyclical indicators and indexes with the cooperation of members of the NBER staff (Zarnowitz and Boschan 1975a and 1975b). Between 1967 and 1975 inflation intensified and two recessions occurred (1970 and 1973–75). Some of the old nominal indicators needed to be deflated to perform well under the new conditions; other revisions became advisable for other reasons, including as usual changes in the available data. The last, rather limited review of the indicators was completed by the BEA staff in 1989.

Table 11.5 presents the complete record of all the revisions in the short lists of the leading indicators for both revivals and recessions from 1950 through 1989. These are also the lists of the components of the successive composite leading indexes since 1960. The numerous footnotes cite the reasons for each of the changes made, as given at the time. They also cover any interim revisions made between the five dates when the new lists were published. Because the table provides detailed references and documentation, the text to follow is limited to a few general explanations.

A quick glance at table 11.5 seems likely to give the impression of great variability, as if almost all selections were transitory and the composition of the lists or indexes was quite unstable over time. But this would be very misleading. A closer look shows that each list retained several series from the previous one and that most of the changes were substitutions or additions within the same economic process groups (see Summary at the end of the table). In other words, most of the revisions replaced some series with other representations of similar variables. Often the reason was the availability of new data deemed conceptually or statistically superior or the discontinuation of old data. Some changes, as noted earlier, were dictated by considerations of timeliness; others, by unpredictable events in government or business (see, e.g., nn. *d* and *k;* the latter had partly to do with the lack of good current data on new business telephones after the breakup of AT&T in 1982). To be sure, at times errors were made and reversed when recognized (see nn. *e* and *cc*).

To the extent that revisions take time but improve the data, there is inevitably a conflict between the promptness and accuracy of estimates and forecasts. But the presumption that the more recent indexes are better does not mean that the earlier ones have failed. On the contrary, as shown in chapter 10, sec. 10.7, the 1950 indexes, though based on pre-1938 data, would have performed reasonably well in terms of timing at the business cycle turns of 1948–75.

11.2.3 Index Construction Methods and Revisions

There are six basic steps in the calculation of a composite index in the traditional NBER-BEA style. (1) Month-to-month percentage changes are computed for each component series expressed originally as levels.[15] Series expressed as changes are differenced. (2) For each series, the changes obtained in step 1 are divided by their long-run average without regard to sign.

15. Formulas ensuring symmetrical treatment of positive and negative changes are used. For a more detailed narrative and algebraic explanation of the methodology, see U.S. Department of Commerce, Bureau of Economic Analysis 1984, pp. 65–69.

Table 11.5 Five Successive Lists of Leading Indicators, 1950–89

Group & Number of Economic Process[a]; Title, Source, & No. of Series[b]	Indicator Lists of				
	1950 (1)	1960 (2)	1966 (3)	1975 (4)	1989 (5)
Marginal Employment Adjustments (I)					
Average workweek, production workers, manufacturing (3; 1)	x	x	x	x	x
Gross accession rate, manufacturing[c]		x			
Nonagricultural placements, all industries[d]			x		
Layoff rate, manufacturing[e]		x		x	
Average weekly initial claims, state unemployment insurance inverted (2; 5)					x
Orders & Deliveries—Consumption & Trade (III)					
New orders, durable goods, industries, current dollars (2; 6)[f]	x	x	x		
New orders, consumer goods & materials, constant dollars (1, 2; 7)				x	x
Change in manufacturers' unfilled orders, durable goods, constant dollars, smoothed (2; 26)[g]					x
Vendor performance—slower deliveries diffusion index (32)[h]				x	x
Index of consumer expectations (83)[i]					x
Formation of Business Enterprises (IV)					
New incorporations, number (13)	x				
Net change in business population, Q[j]		x			
Index of net business formation (1; 12)[k]			x	x	
Business Investment Commitments (IV)					
Commercial & industrial building contracts, floorspace (9)[l]	x	x			
Contracts & orders for plant & equipment, current dollars[m]				x	
Contracts and orders for plant and equipment, constant dollars (1, 2; 20)[n]				x	x
Residential Construction Commitments (IV)					
Residential building contracts, floorspace[o]	x				
Housing starts, new private units (2; 28)		x			
Building permits, new private housing units (2; 29)[p]				x	x
Inventory Investment (V)					
Change in business inventories, current dollars, Q (1; 245)[q]		x			
Change in book value, manufacturing & trade inventories (1, 2; 31)[r]			x		
Net change in inventory on hand & on order, constant dollars smoothed (1, 2; 31)[s]				x	

Table 11.5 Continued

Group & Number of Economic Process[a]; Title, Source, & No. of Series[b]	Indicator Lists of 1950 (1)	1960 (2)	1966 (3)	1975 (4)	1989 (5)
Sensitive Commodity Prices (VI)					
Wholesale price index, 28 basic commodities[t]	x				
Industrial materials price index[u]		x	x		
Percentage change, WPI of crude materials, excluding foods & feeds, smoothed (1, 3; 99)[v]				x	
Percentage change in a revised index of sensitive materials prices (1, 3; 99)					x
Stock Prices (VI)					
Dow-Jones index, industrial stock prices[w]	x				
S&P stock price index, 500 common stocks (19)[x]		x	x	x	x
Profits and Profit Margins (VI)					
Corporate profits after taxes, current dollars, *Q* (1; 16)[y]	x	x			
Ratio, price to unit labor cost, manufacturing (1, 3; 26)[z]		x			
Credit Difficulties (VII)					
Current liabilities of business failures (14)[aa]	x	x			
Credit Flows (III)					
Net change in consumer installment credit (4; 113)[bb]		x			
Money (VII)					
Money supply, M1, in constant dollars (1, 4; 105)[cc]				x	
Money supply, M2, in constant dollars (1, 4; 106)[cc]					x
Percentage change in total liquid assets, smoothed (1, 4; 104)[dd]				x	
Summary Number of:					
Component series in each list	8	12	12	12	11
Series unchanged from previous list	n.a.[ee]	4	5	4	7
Substitutions within same group	n.a.	4	4	5	2
Additions within same group	n.a.	2	2	1	2
Additions in new group	n.a.	2	1	2	0
Deletions from previous list		4	3	3	2

Sources: Moore 1950, 1961; Moore and Shiskin 1967; Zarnowitz and Boschan 1975a; Hertzberg and Beckman 1989. See text for more detail.

[a]Roman numerals denote economic processes. See table 11.1 for the titles.

[b]The first number in parentheses following the title of the series denotes the source as follows: 1. U.S. Department of Commerce, Bureau of Economic Analysis; 2. U.S. Department of Commerce, Bureau of the Census; 3. U.S. Department of Labor, Bureau of Labor Statistics; 4. Board of Governors of the

(*continued*)

Table 11.5 Continued

Federal Reserve System. The second number in parentheses identifies the series in the *BCD* and the *Survey of Current Business* (*SCB*). The numbers are given only for the series carried most recently in the *BCD* or currently in the *SCB*. For the corresponding information on the other series, see footnotes below.

[c]Source 3; old *BCD* series no. 2. Replaced in 1966 by nonagricultural placements (source 3; old *BCD* series no. 4), which "provide broader coverage and prompter availability" (Moore and Shiskin 1967, p. 69; henceforth M-S). Discontinued in February 1982.

[d]Replaced in September 1969 by average weekly initial claims for state unemployment insurance (see note *e* below). Stated reason: Shift in emphasis of the public offices of the U.S. Department of Labor from total placements to services for disadvantaged workers. These services take more staff time than is required for qualified workers. The effect of this policy change has been to slow down the number of placements. Data not updated since June 1970, discontinued July 1971 (see *BCD*, September 1969 and July 1971, p. III).

[e]Source 3; old *BCD* series no. 3. Omitted in 1966 because "workweek and placements enough for short list" (M-S). Reinstated to replace average weekly initial claims in 1975 because "layoff rate leads more consistently at troughs" (Zarnowitz and Boschan 1975a, p. 6; henceforth Z-B). Discontinued in February 1982 and replaced again with average weekly initial claims.

[f]Replaced in 1975 by a series of real new orders received by manufacturing industries producing primarily consumer goods and materials. This allows separation of these orders from orders for equipment, which are aggregated with contracts for plant (see note below). Deflation needed for better cyclical performance since the late 1960s because of persistent rise in the general price level (Z-B).

[g]Includes the on-order portion of the series net change in inventories on hand and on order in constant dollars, which was dropped from the leading index in 1989 because the current data for the on-hand portion are not available in time to be included in the initial estimate of the index. Covers all durable goods, including also capital goods and defense products. (See Hertzberg and Beckman 1989, pp. 97–98; henceforth H-B).

[h]Based on surveys by Purchasing Management of Chicago and, since 1989, by the National Association of Purchasing Management. "Best available indicator of changes in delivery lags" (Z-B).

[i]Compiled by the University of Michigan's Survey Research Center. "Added to provide a new dimension to the leading index" (H-B).

[j]Source: Office of Business Economics, U.S. Department of Commerce. Quarterly. Replaced the monthly Dun and Bradstreet, Inc., series on new incorporations in 1960 because the latter "is occasionally affected strongly by changes in legislation, especially in the tax laws, . . . and partly because the net changes take account of discontinuances" (Moore 1961, p. 66; henceforth M61). Discontinued after 1959:4 and replaced by the monthly index of net business formation. For references and comparisons, see M61, ch. 14).

[k]Original sources: Dun and Bradstreet, Inc., and Bureau of the Census. Suspended from the index in March 1987. This series "deteriorated as a measure of change in business population, primarily because of the poor quality of one component of the series and the unavailability of data of another component in the time for inclusion in the initial release" (see *BCD*, February 1987, p. III).

[l]Source: McGraw-Hill Information Systems Company. Replaced in 1966 by contracts and orders for plant and equipment (see note *m* below) because of "poor timing and conformity record since 1950. The equivalent value series, plus privately owned public utilities, is included in the series on contracts and orders" (M-S).

[m]"Most comprehensive series on new investment commitments by business enterprises"; new since 1960 list (M-S).

[n]"Deflation is needed for better cyclical performance since the late 1960's" (Z-B).

[o]Source: F. W. Dodge Corp. Replaced by housing starts in 1960 because starts are "less erratic . . . and currently published more promptly" (M61, p. 65).

[p]Compared with housing starts, "permits series is smoother" (M-S).

[q]Volatile but persistently leading, especially at peaks, since 1939 (M61, pp. 68–69).

[r]Replaced the quarterly inventory investment in 1966 because "monthly series more current" (M-S).

[s]"Concept of including stocks on order is better. Deflation is needed for better cyclical performance" (Z-B).

Table 11.5 Continued

[t]Source: BLS. Unweighted geometric mean of relatives for 28 products.

[u]Source: BLS. Includes daily prices of 13 raw or simply processed materials. Conformed more closely to business cycles than the broader index. which includes prices of foodstuffs "subject to vagaries of weather conditions and government farm price policies" (M61, p. 69).

[v]"Percent change is better than level. Leads are more consistent. especially since the late 1960's (Z-B). Revised in 1989 to improve the methodology and date consistency (H-B).

[w]Source: Dow-Jones & Co., Inc. Unweighted mean of prices of 30 industrial stocks.

[x]Source: Standard & Poor's Corp. New monthly index of prices of 500 common stocks introduced in 1957; before then coverage smaller but increasing (from 198 stocks in 1918 to 480 in 1957). The S&P index has the advantage of greater coverage and diversification compared with the Dow-Jones index.

[y]Dropped from the index in 1975 because of being "quarterly and tardy (low score for currency)" (Z-B).

[z]Old *BCD* series no. 17. Ratio of price to unit cost, manufacturing. dropped from the index in 1975 because it "failed to lead at the last three business cycle troughs (1958–70)" (Z-B).

[aa]Source: Dun & Bradstreet, Inc. Dropped from the index in 1966 because of "poor timing and conformity record since 1948" (M-S).

[bb]Included in the index in 1966 because of "wide cyclical movements and consistent leads" (M-S). Dropped from the index in 1975 with the comment that it "lacks timeliness. In recent period. very erratic and more nearly coincident than leading at troughs" (Z-B).

[cc]Deflated by the consumer price index. Included in the index in 1975 as an "important measure of the quantity of money in real terms. Good scores for indicator performance" (Z-B). Replaced in March 1979 by M2 money supply deflated by the CPI. The more comprehensive monetary aggregate was favored by changes due to deregulation of banking, new interest-earning and highly liquid types of deposits. and greater incentives to economize on the use of money.

[dd]Included in the index in 1975 as a "comprehensive measure of changes in wealth held in liquid form by private nonfinancial investors" (Z-B). Replaced in February 1983 by change in business and consumer credit outstanding (1. 4: 111). which had a better record of cyclical conformity and timing in recent years. The credit change series was dropped from the index in January 1989.

[ee]n.a. = not applicable.

This standardization makes the mean absolute value for each so-transformed series equal to 1, and it prevents the more volatile series from dominating the index. (3) For each month, the values of all available component series obtained in step 2 are averaged.[16] (4) The resulting changes are then adjusted for the leading (L) and lagging (Lg) indexes so that their long-run absolute averages are equal to that of the coincident index (C). This index standardization procedure (division by the ratio of long-run averages L/C and Lg/C, respectively) has the purpose of facilitating the use of the three indexes as a consistent system. (5) The so-modified average changes are cumulated into a "raw" index for each of the three sets of indicators. (6) A common trend is established for the three indexes, which can be viewed as a linear approximation to the secular movement in aggregate economic activity.[17] The trend adjustment

16. Before the last BEA revision (i.e., through January 1989) the average was a weighted one, with the total scores of the series serving as weights. Since then. in the computation of the revised indexes. the components of each index are assigned equal weights. The change had relatively little effect because the weights based on the scores of the series were not very different from 1 (see Hertzberg and Beckman 1989. p. 99).

17. Before 1989 the target trend was the average of the long trends in the four components of the coincident index (0.268% per month). For the latest revised indexes, the trend is that of real GNP (0.261% per month). Clearly. the two trends are almost the same. so the effect of this change was minimal (Hertzberg and Beckman 1989. p. 100).

enhances the usefulness of the indexes by making them differ only with respect to short-term, mainly cyclical movements.

In three decades of development, the composite indexes of leading, coincident, and lagging indicators spread to many countries, aided by and promoting the growth of information about national and international economic fluctuations (Klein and Moore 1985; Moore and Moore 1985). Following Mintz 1969, methods of combining series in terms of deviations from trend or rates of growth were devised to track growth cycles, which have become more common than business cycles. These methods share the several basic elements of index construction listed above, but they adapt them depending on the treatment of trends and weighting and differ in many technical details. Interesting work on composite index construction has been done in Australia (Haywood 1973) and by the British Central Statistical Office (CSO 1976), Statistics Canada (Rhoades 1982), the Organization of Economic Cooperation and Development (OECD 1987), and Japan's Economic Planning Agency (EPA, since the early 1970s). A useful comparison of these methods is provided in Boschan and Banerji 1990, where a procedure is suggested that uses the standard deviation of smoothed, detrended series as the standardization factor. The amplitude standardization is at the core of the construction of these indexes. It is itself a form of weighting that equalizes the volatility of the index components. Scoring is useful in choosing the component series but weighting based on the scores may or may not be helpful (it could be redundant). Transformations of the individual series are needed to induce stationarity as appropriate in each particular case.

The U.S. composite indexes are updated near the end of each month by computing the preliminary figures for the previous month and recomputing a number of preceding values.[18] From time to time, recomputations are made to incorporate longer term, benchmark and seasonal revisions. In addition, the standardization factors for the components and the indexes, and the trend adjustment factors, are updated at longer intervals.[19]

These statistical revisions, as well as the definitional revisions (compositional and weight changes), make it necessary to differentiate clearly between the uses of preliminary estimates and the revised index data. Evaluations of the latter have their role in historical and analytical contexts, but authentic ex ante or real-time forecasting always involves preliminary values. Diebold and Rudebusch (1987) examined the stochastic properties of the first-released data and subsequent revisions for the Department of Commerce index of leading indicators. They report that in 1983–87 the revisions (particularly the earliest ones, which contained the most information) behaved approximately as if they were errors of efficient forecasts. This was not the case in 1979–81, when measurement errors due to the missing index components were much more severe.

18. For 11 months before 1987 and for 5 months in the latest revised indexes.
19. Eight times between August 1970 and January 1989.

11.3 Forecasting with Leading Indicators

11.3.1 Some General Observations

By their own testimony, practicing forecasters use leading indicators widely, along with other methods and information. In predicting the course of the economy in the near future, business economists generally rely on "a complicated combination of quantitative and qualitative elements—with a heavy emphasis placed upon the exercise of judgment" (Butler, Kavesh, and Platt 1974, p. 207). The major expenditure components of GNP are forecast with the aid of theory, statistical procedures such as regression, and assumptions about policy and other external changes. The GNP accounting framework and checks and adjustments for consistency and plausibility are then used to derive the overall forecasts. As documented in chapter 13 (sec. 13.8), evidence from surveys shows that most professional analysts of the short-term economic outlook favor this flexible and eclectic approach, labeled the "informal GNP model." When asked to rank items on a short list of general forecasting techniques according to their own usage, most respondents chose the informal GNP model as first. But very few replies referred only to this or any other single method; combinations prevail heavily. Many of the surveyed forecasters reported using leading indicators and outside macroeconometric models, which were given predominantly second or third ranks. (Of course, some forecasters work primarily with their own full-scale econometric models, but their number is relatively small.) Anticipations surveys and "other methods" (such as time-series analysis) received fewer references and were cast mostly in subsidiary roles.

The methods are broadly defined and jointly applied so that it is hardly possible to isolate and assess the contribution of each (more on this in chapter 13). The surveys provide no measures of just how well the leading and confirming indicators and indexes serve their users, although they affirm credibly that these time series are in fact regularly and extensively employed as tools for the analysis of current and forecasting of future business conditions.

To test the predictive value of the leading indicators, it is necessary to specify the data (usually a leading index and/or its components); the target (all changes in some representation of aggregate economic activity; turning points in business cycles or the selected aggregate); and the procedure or rule to be applied (regressions of coincident on leading variables; filtering of the leading series to forecast peaks and troughs). A double dichotomy has emerged in the literature. (1) Some writers concentrate on the prediction of turning points, arguing that this is the most appropriate or the most important function of leading indicators, and that the economy behaves differently in the vicinity of peaks and troughs or differently in expansions than in contractions. Others use the leaders simply as linear predictors of all successive values of some measure of economic activity. (2) Most tests of either sort have been based on revised data as of the time of the study, but some have tried to re-create the actual, real-time forecasting situation by employing preliminary data available at the time of the forecast.

11.3.2 Tests of Turning-Point Predictions

Several studies attempted to evaluate leading indicators and indexes as predictors of cyclical turns, that is, sustained changes of direction in aggregate economic activity. Such appraisals require that a set of filtering rules be considered. Not each monthly wiggle in the leading index need be significant, and excessive concern with a single month's movement must be severely discouraged.

Vaccara and Zarnowitz (1978a, 1978b) considered two very simple rules. (1) Let three consecutive monthly declines in the leading index (1975 version) signal a downturn, and three consecutive rises, an upturn. This yielded mean leads of 8.8 and 3.2 months at business cycle peaks and troughs in 1948–76, no missed turns, and five "extra cycles."[20] Because 3 months must elapse before a turning point is identified, a minimum lead of 4 months is required for a sufficiently early warning. Leads at least this long did occur at all five peaks covered but only at three of the six troughs. (2) As an alternative, treat each directional change in the trailing 6-month moving average of the index as a turning-point signal. This rule reduces the requisite effective lead time to 2 months. The mean leads are here 5.6 and 1.0 months at peaks and troughs, respectively. Again, there are no missed turns and five false signals each of peaks and troughs.

Earlier, Hymans (1973) applied the rule of two consecutive monthly declines (rises) in expansion (contraction) to the leading index constructed in 1966. As might have been expected, this resulted in an unacceptably large number of false predictions, particularly of peaks. The assumption that the forecaster has timely knowledge of the current phase of the business cycle is hardly justified in the critical periods around peaks and troughs. Hymans constructed an alternative index of leading indicators based on spectral analysis, using the same components as the 1966 NBER index. This method combines indicator components corresponding to four cyclical periodicities (of 60, 40, 30, and 24 months) and so filters out false signals much more effectively, but at the cost of strongly reducing the size and frequency of the leads.

Filtering rules typically involve trade-offs of currency for accuracy: the more smoothing, the fewer the extra turns but also the shorter the effective lead time at true turns. This is so whether one requires longer runs (more months of maintained direction of movement) or longer spans (more months over which to measure the change). To get better signals, the magnitude of changes in the indicators should be taken into account as well as their direction.

20. Three extra cycles were associated with the growth cycle slowdowns in 1950–51, 1962, and 1966. The other two involved declines that lasted only 3 months each.

Combining selected leading series into indexes can greatly reduce the frequency of false predictions of turning points but cannot eliminate the problem. The leading indexes issue warnings not only of recessions but also of cyclical retardations, that is, phases when the overall real growth rates fall below their long-term average but remain positive (except possibly for an occasional brief slippage below 0). In sum, these indexes predict best the "growth cycles" in trend-adjusted aggregates of output and employment. This was found to be true as well for Japan, Canada, and Western Europe (Moore 1983, chs. 5 and 6). What is needed, therefore, is a method that would allow us to discriminate in a timely fashion between recessions and major slowdowns and thereby safeguard against the main type of "extra turns" that occur in practical forecasting.

To this purpose, a sequential signaling system has been devised and tested with some promising results in Zarnowitz and Moore 1982. The long-term trends built into both the leading and the coincident indexes are 3.3% per annum, and the standard deviations of their random components are approximately 1.0%. Now consider smoothed 6-month percentage changes in these indexes, calculated at compound annual rates from preliminary data available at the time of forecast.[21] This yields two trendless series to be labeled the "leading index rate" (L) and the "coincident index rate" (C). The procedure is to monitor their movements on a current basis, paying special attention to two percentage bands: 3.3 ± 1.0 and 0 ± 1.0. The first signal of a business cycle peak (P1) occurs when the leading index rate falls below the upper band, while the coincident index rate, which will typically be higher, remains nonnegative. In short, P1: $L < 2.3$, $C \geq 0$. The second signal of a peak is observed when the leading rate falls below the lower band and the coincident rate falls below the upper band; that is, P2: $L < -1.0$, $C < 2.3$. The third and last signal (P3) is defined by $L < 0$, $C < -1.0$; that is, the coincident rate falls below the lower band, and the leading rate remains negative. A signal is invalidated when either index rate, after declining across a band, rises again above it; it is not invalidated when L or C merely backs up into the band and stays there.

Table 11.6 presents the record of these signals based on historical data for the period through September 1976 and on preliminary data thereafter. The four slowdowns not followed by recessions (in 1951, 1962, 1966, and 1984) were all associated with prompt P1 signals. In 1951, P2 followed P1 by 4 months, adding to uncertainty, but in the other instances there were no further warnings. The absence of P2 should have ruled out a false prediction, and the absence of P3 a false identification, of a business cycle peak. That the signals would have been reasonably timely on most occasions is suggested by the following overall timing means (standard deviations): P1 led P2 by 5 (2)

21. This is done by dividing the current month's index by the average of the 12 preceding months and raising the ratio to the 12/6.5 power (the center of the average is located 6.5 months before the current month).

Table 11.6 Sequential Signals of Slowdown and Recession, 1951–90

Dates of Signals[a]			Business Cycle Peak[b]	Leads (−) or Lags (+), in Months, at Business Cycle Peaks			Growth Cycle Peak[b]	Leads (−) or Lags (+) in Months, at Growth Cycle Peaks		
P1 ($L < 2.3$, $C \geq 0$)	P2 ($L < -1.0$, $C < 2.3$)	P3 ($L < 0$, $C < -1.0$)		P1	P2	P3		P1	P2	P3
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
3/51	7/51	. . .	none	. . .	. . .	. . .	3/51	0	+4	
6/53	8/53	9/53	7/53	−1	+1	+2	3/53	+3	+5	+6
1/56	7/56	9/57	8/57	−19	−13	+1	2/57	−13	−7	+7
9/59	6/60	9/60	4/60	−7	+2	+5	2/60	−5	+4	+7
5/62	. . .	. . .	none	. . .	. . .	. . .	5/62	0	. . .	. . .
6/66	. . .	. . .	none	. . .	. . .	. . .	6/66	0	. . .	. . .
6/69	11/69	4/70	12/69	−6	−1	+4	3/69	+3	+8	+13
8/73	1/74	3/74	11/73	−3	+2	+4	3/73	+5	+10	+12
11/78	5/79	3/80	1/80	−14	−8	+2	12/78	−1	+5	+15
6/81	8/81	10/81	7/81	−1	+1	+3	none	. . .	. . .	. . .
7/84	. . .	. . .	none	. . .	. . .	. . .	6/84	+1	. . .	. . .
11/87	8/90[c]	10/90	7/90	[d]	+1	+3	2/89	−15	+18	+20
Averages: sample period (before 10/1976)										
Mean lead (−) or lag (+)				−7	−2	+3		−1	+4	+9
Standard deviation				7	6	2		6	6	3
Averages: postsample period (since 10/1976)										
Mean lead (−) or lag (+)				−8	−2	+3		−5	+12	+18
Standard deviation				9	5	1		9	9	4

Source: Zarnowitz and Moore 1982; reprinted in slightly updated form in Moore 1983, table 4-7.

[a]For full definition of the signals, see text. Revised data for the indexes (taken from U.S. Department of Commerce, Bureau of Economic Analysis 1977) are used prior to October 1976; preliminary data (taken from various issues of *BCD* through March 1990 and the Survey of Current Business for April 1990 to July 1990) are used for the period October 1976 to June 1990.

[b]Dated according to the NBER reference chronologies.

[c]An earlier P2 signal (2/90) was canceled in 5/90 when $L > 1.0$.

[d]The early 11/87 signal cannot be meaningfully related to the business cycle peak of July 1990.

months, P2 led P3 by 5 (5) months, and growth cycle peaks led business cycle peaks by 7 (4) months.

Sequences of all three signals occurred before each of the seven recessions covered, and only at these times.[22] They were short and concentrated when there was little or no warning from a prior slowdown, as in 1953 and 1981; they were long and spread out when protracted slack phases preceded the downturns. If very early signals are "premature" in the sense of lacking credibility, P1 and P2 in 1956 and P1 in 1978 would qualify as such because each of these events anticipated the corresponding business cycle peaks by more than a year.

The "band approach" to sequential signaling produced no false alarms at all during the turbulent years 1976–83 (the first half of our postsample period).[23] This success is probably attributable in part to the fact that no slowdown-without-recession occurred in that period. But the P1 signals also announced and accompanied the 1984–86 retardation in its first 1.5 years; then T2 and T3 signals predicted the end of this phase in its last year (table 11.7).[24] Another P1 signal appeared first in November 1987, after the stock market crash, but it was neither canceled nor confirmed by P2 before the new slowdown, which is now known to have started in or about February 1989. A P2 warning flashed in February 1990 but was canceled in May and reinstated in August; and a P3 signal confirmed the downturn in October.

At business cycle troughs, the first signal (T1) occurs when the L rate rises across the lower band (0 ± 1%) while the C rate remains less than 1.0%. The second signal (T2) is given when L rises across the upper band (3.3 ± 1%) and C rises above 1.0%. The third signal (T3) requires that both L and C cross the upper band from below, that is, rise above 4.3%. False signals of troughs are defined, analogously to those of peaks, as reverse crossing through the bands. They are infrequent and therefore relatively unimportant. As demonstrated in table 11.7, the signals of business cycle troughs are less dispersed but much tardier than those of peaks. Full recognition of peaks (P3) required 1–5 months after the event, not long considering the irreducible lags of measurement and detection in noisy data. In contrast, full recognition of troughs (T3) required 5–13 months (cf. col. 7 in tables 11.6 and 11.7). The longer lags at troughs simply reflect the lack of earlier signals from the indicators.

22. Sufficient data are not available to make the calculations for the 1948 peak, but P3 arrived no later than January 1949 and possibly earlier.

23. The reverse crossings through the bands, which constitute the false signals in this approach, are very infrequent for both index rates, L and C, whether preliminary or first revised data are used. Also, the percentages of the observations falling into the two bands have been relatively small for L and C in 1948–81 (but larger in the 1980s). In contrast, spurious turning-point predictions abound in the alternative "level approach" when currently available preliminary data are used. In this system, critical levels of 3.3% and 0% perform the functions of the bands 2.3%–4.3% and −1.0%–+1.0%, respectively. When the first revised data are used, which implies a delay of 1 month, the frequency of false signals is greatly reduced. See Zarnowitz and Moore 1982.

24. In growth cycles where no P3 signal is reached, T1 is usually skipped (but see table 11.7, n. *c*). T2 then may cancel a previous P1 or P2 signal; T3 marks renewed above-average growth.

Table 11.7 **Sequential Signals of Recovery, 1949–86**

Dates of Signals[a]			Business Cycle Trough[b]	Leads (−) or Lags (+), in Months, at Business Cycle Troughs			Growth Cycle Trough[b]	Leads (−) or Lags (+), in Months, at Growth Cycle Troughs		
T1 ($L > 1.0$, $C < 1.0$) (1)	T2 ($L > 4.3$, $C > 1.0$) (2)	T3 ($L > 4.3$, $C > 4.3$) (3)	(4)	T1 (5)	T2 (6)	T3 (7)	(8)	T1 (9)	T2 (10)	T2 (11)
8/49	1/50	3/50	10/49	−2	+3	+5	10/49	−2	+3	+5
1/52[c]	3/52[d]	8/52	none	. . .	. . .	. . .	7/52	−6	−4	+1
5/54	11/54	12/54	5/54	0	+6	+7	8/54	−3	+3	+4
6/58	10/58	11/58	4/58	+2	+6	+7	4/58	+2	+6	+7
3/61	6/61	8/61	2/61	+1	+4	+6	2/61	+1	+4	+6
. . .	2/63	5/63	none	. . .	. . .	. . .	10/64	. . .	−20	−17
. . .	5/67	11/67	none	. . .	. . .	. . .	10/67	. . .	−5	+1
11/70	5/71	12/71	11/70	0	+6	+13	11/70	0	+6	+13
6/75	9.75	11/75	3/75	+3	+6	+8	3/75	+3	+6	+8
9/80	12/80	4/81	7/80	+2	+5	+9	7/80	+2	+5	+9
7/82	4/83	6/83	11/82	−4	+5	+7	12/82	−5	+4	+6
. . .	12/85	4/86	none	. . .	. . .	. . .	1/87	. . .	−13	−9
Averages: sample period (before 10/1976)										
Mean Lead (−) or Lag (+)				+1	+5	+8		−1	0	+3
Standard Deviation				2	1	3		3	9	8
Averages: postsample period (since 10/1976)										
Mean Lead (−) or Lag (+)				−1	+5	+8		−2	−1	+2
Standard Deviation				4	0	1		5	10	10

Source: Zarnowitz and Moore 1982, reprinted in slightly updated form in Moore 1983, table 4-8.

[a]See n. *a* in table 11.6.

[b]See n. *b* in table 11.6.

[c]Note that the C rate fell below 1% but not below 0 in the slowdown of 1951–52; that is, no P3 occurred before this T1 signal.

[d]The C rate fell below 0 for a single month in July 1952, probably in connection with the major steel strike, but both L and C rates rose sharply beginning in August.

As for the *average* timing at business cycle peaks, P1 led by 7–8 months in both the sample and the postsample periods, P2 led by 2–4 months, and P3 lagged by 2–3 months. Shorter leads and longer lags characterized the timing of the same events at growth cycle peaks. T1 tended to signal business cycle troughs at approximately the time of their occurrence, T2 5 months later, and T3 8 months later. At growth cycle troughs, the leads and lags of the T signals were on average shorter but much more dispersed.

In terms of the latest revised data first published by the Department of Commerce in March 1989, the means (standard deviations) of the leads or lags of the signals, in months, are as follows:

	P1	P2	P3	T1	T2	T3
At business cycle turns	−7 (6)	−3 (5)	+2 (1)	0 (2)	+5 (1)	+8 (3)
At growth cycle turns	−2 (6)	+4 (5)	+9 (3)	−1 (2)	−1 (9)	+4 (8)

These statistics are closely similar to those computed from earlier and preliminary data in tables 11.6 and 11.7 (all discrepancies fall in the range of −1 to +1 months). To be sure, comparisons of individual observations disclose some larger effects of the revisions, but even so, few of the differences would matter greatly.

11.3.3 Probabilities of Recession and Recovery

Several recent papers discuss the implications of the assumption that macroeconomic variables are subject to two different probability distribution functions, one of which applies in business expansions and the other in contractions. The switches from the former to the latter or vice versa are supposed to occur suddenly at random time points and be unobserved; they can only be inferred from the data according to some model and prediction rules. Thus Neftçi (1982) splits the data on the composite index of leading indicators for 1948–70 into "downturn regimes" and "upturn regimes," apparently relying in the main on judgment. He then smoothes the historical frequency distributions of the monthly changes in the index with a 3-month centered moving average to estimate the probability distributions separately for the two regimes. His formula for assessing the probability of recession is recursive and dynamic in that it includes the previous month's outcome and cumulates the probabilities from 0 at the start of each expansion to 100% at the end. It also involves a prior probability distribution based on the assumption that the likelihood of a downturn increases slightly in each month as the expansion ages.[25]

25. Let p_t be the probability of an imminent peak this month (p_{t-1}, last month); let p_t^e be the probability that the latest observation came from the distribution for expansions, and p_t^c that it came from the distribution for contractions; and let Π_t be the prior probability of a peak based solely on the length of expansion to date. Then

$$p^t = \frac{[p_{t-1} + \Pi_t (1 - p_{t-1})p_t^c}{[p_{t-1} + \Pi_t(1 - p_{t-1})p_t^c] + [(1 - p_{t-1})p_t^e(1 - \Pi_t)]} .$$

A signal of a nearby recession is issued when the probability computed by the formula exceeds a preset level of confidence. Neftçi, in two illustrations using the 90% level, shows that the leading index would have given warnings of recession in August 1973 and April 1979 (1982, pp. 238–40).[26]

The assumption that the life expectancy of an expansion is a steadily declining function of its duration is questionable (see chapter 8, sec. 8.3.5) and so is Neftçi's judgmental dating, which deviates considerably from the NBER chronologies. The distinction between business cycles and growth cycles needs to be made explicit in this context. The target dates are never clearly defined, which makes it difficult to assess the proposed method and verify its results (the definition of a recession as two consecutive declines in GNP [see 1982, p. 232] is simply inadequate). But these defects of implementation do not detract from the importance of Neftçi's basic idea that prediction of turning points can be treated as an optimal stopping-time problem (Shiryayev 1978). The concept stimulated much interesting work on how the statistical technique of sequential analysis can best be used to determine when the cumulative information justifies announcing a turning point, given the cost of a false signal and a preset margin of error.

Applying this analysis to the performance of the leading index at post-1948 business cycle peaks, Palash and Radecki (1985) report the following estimates, in months: 7/53, −1; 8/57, −14; 4/60, −5; 12/69, −1; 11/73, −3; 1/80, −6. These lead times do not differ much from those of the P1 signals in table 11.6, column 5, but the absence here of any prior warning of the 1981 downturn is troubling. Also, the probability approach produces signals of slowdowns as well as recessions but, unlike the band approach, does not attempt to distinguish between the two by means of additional information.[27]

Diebold and Rudebusch 1989 is a comprehensive evaluation of cyclical turning-point forecasts from historical data on the BEA composite index of leading indicators for December 1949 to December 1986. The method is a Bayesian sequential probability recursion (SPR), which retains some aspects of Neftçi's approach and modifies others.[28] Here all probabilities are used directly as forecasts, and the stopping rule is not applied. A measure of *accuracy* for a set of n such forecasts is the quadratic probability score

26. Note the closeness of these dates to their counterparts in table 11.6, cols. 1 and 2 (a P1 signal in 8/73 and a P2 signal in 5/79).

27. Thus the results of Palash and Radecki include continuous predictions of recession from 4/51 to 5/52 (the slow-growth phase of the Korean War cycle; cf. table 11.6, line 1). They also show a 6-month lead at a misnamed "business cycle peak" of 12/66, which corresponds to the P1 signal of 0 at the growth cycle peak of 6/6 in table 11.6.

28. Among the retained elements is the assumption that the composite leading index (CLI) has the following simple probability structure: $\Delta CLI = \alpha^u + \varepsilon_t^u$ in expansions; $\Delta CLI = \alpha^d + \varepsilon_t^d$ in contractions (where the α's are constants and the ε's are independently and identically distributed random variables with means of 0). The densities corresponding to the two regimes, f^u and f^d, are estimated by fitting normal distribution functions to the CLI data partitioned according to the specific-cycle peaks and troughs in the index that match the NBER business cycle turns. The probability of a turning point is taken to be independent of the duration of the phase; hence the transition priors are constant.

$$QPS = \frac{1}{n}\sum_{t=1}^{n} 2(p_t - r_t)^2,$$

where r_t is the realization corresponding to the predicted probability, p_t (it equals 1 when a turning point has occurred and 0 otherwise). QPS has a range from 0 (best) to 2 (worst). A measure of *calibration* (i.e., closeness of predicted probabilities to observed relative frequencies) is the global squared bias, GSB = $2(\bar{p} - \bar{r})^2$, where $\bar{p}$ and $\bar{r}$ are means of p_t and r_t, respectively. GSB, too, ranges from 0 to 2. Forecasts are also scored on other criteria, notably their local discriminating power, or *resolution,* which reflects the extent to which different probability values are followed by different realizations. All scores are for several forecast horizons (1, 3, 5, 7, 9, and 13 months). Only symmetric loss functions are considered.[29] However, the performance of the leading indicators tends to be asymmetric and such that the index would presumably score better if missing a true turn mattered more than calling a false turn.

The SPR forecast scores on accuracy (QPS) at peaks increase from .29 for the shortest to .49 for the longest horizon. Surprisingly, these results are worse than those obtained from the simple rule that yields probabilities of 1 when the leading index shows three consecutive declines (3CD) and of 0 otherwise (the scores for these 3CD forecasts increase from .19 to .34 in the 1–9 months range). Even the naive forecast of zero probability of a downturn throughout earns lower (i.e., better) scores than SPR. The optimal constant-probability forecast, where the constant is chosen to minimize QPS for expansions, performs much better than any of these methods. However, constant priors can be found that optimize the posterior probability predictions via SPR, and this method (SPR*) produces about equally good results (scores of .04–.39). Also, both SPR and SPR* show higher relative accuracy when the LPS measures are substituted for QPS (see n. 29), whereas 3CD performs poorly here.

At troughs, SPR* generally scores best; the naive forecast worst. The optimal constant-probability forecast for contractions comes in a close second; SPR third. The scores generally increase with the horizon in the range of 1–7 months (e.g., from .27 to .48 for SPR under QPS).[30]

The SPR predictions tend to overpredict the probabilities of turning points; the naive no-change predictions obviously underpredict them. All other examined types of forecast, from the simple 3CD rule to the sophisticated SPR*, are well-calibrated or unbiased (i.e., have GSB $\approx$ 0 for all horizons). The resolution scores show SPR and SPR* to be moderately and about equally informative, whereas the naive and constant-probability forecasts are, of course, by their very nature of no value in this respect.

29. They are implied by both QPS and an alternative, the log probability score (LPS). The LPS penalizes large errors more heavily than QPS does, and so the absolute level of the scores tends to be higher for LPS than for QPS (cf. tables 1 and 2 in Diebold and Rudebusch 1989, p. 382).

30. The longer predictions show improvements that are correctly recognized as artificial. Because the postwar recessions were short (all but one less than a year), it was easy to predict that they would end after 9 or 11 months.

In absolute terms, all predictive techniques performed worse in contractions (at troughs) than in expansions (at peaks). In relative terms, the SPR method performed best in contractions. The longer their horizon (time span to target), the less accurate would the probabilistic turning-point predictions typically be. This is a general property of optimal forecasts and a widely observed regularity, as documented in part IV of this book.

The Diebold and Rudebusch 1989 analysis is ex post in the sense that their probability densities are estimated over the entire sample and their data on the leading index represent the latest "final" revisions. An ex ante evaluation based on rolling probability densities, preliminary and partially revised data, and the SPR method (Diebold and Rudebusch 1991) yields closer approximations to real-time out-of-sample forecasts. Not surprisingly, the so-constructed ex ante forecasts are less accurate than their ex post counterparts. The main reason is probably data errors rather than pure forecast errors. As we have seen, not only are the component series of the leading index subject to frequent short-term statistical revisions but they have not all been available in time for inclusion in the first preliminary values of the index; moreover, the composition of the index itself was altered on several occasions. The probability forecasts based on preliminary data reflect all these measurement errors and could be very sensitive to the compositional changes in particular. But the NBER business cycle turning-point dates (the target of these forecasts) are established retrospectively, in light of historical, revised information. Many important macroeconomic series (e.g., real GNP) undergo repeated and often relatively large revisions, and forecasters on the whole predict the future values of these variables better in preliminary than in revised form.[31]

Hamilton (1989) applies a nonlinear iterative filter to quarterly growth rates in real GNP for 1952–84 and reports the maximum likelihood estimates of the parameters of the underlying process. He notes that one possible outcome of this use of Markov switching regression to infer regime changes might have been an identification of above-trend ("fast") and below-trend ("slow") growth phases. However, this was not the case; instead, the best fit to the data separated periods of positive growth from periods of negative growth (recession). Indeed, the dates of the switches determined by this method agree quite well with the NBER business cycle chronology.[32]

Hamilton's statistical analysis extends that of Neftçi. That its results agree so well with the results of the historical NBER studies is remarkable. But here again it is well to remember that the focus of all of this work is on the U.S.

31. For evidence, see Zarnowitz 1967 and R. Cole 1969a; for more on the general properties of macroeconomic forecasts, see part IV below. True ex ante predictions must use preliminary data for the most recent periods, and the effective weight of this information will often be large. Thus the forecasts are apt to have more in common with the early than with the later revised values of the target variables.

32. The only large discrepancies are associated with Hamilton's peak dates of 1957:1 and 1979:2, which coincide with oil price increases (the NBER dates are 1957:3 and 1980:1). Elsewhere the differences are mostly zero; in a few cases, ± 1 quarter.

business cycle chronology derived from ex post information. Slowdowns are more frequent than recessions, the two are not so easily distinguishable on a current basis, and leading indicators are sensitive to both. Growth cycles attract much popular and official attention, particularly abroad, and the concept receives support from some contemporary theories. The leading and confirming indicators in Europe and Japan are mostly geared to fluctuations in deviations from trend.

Niemira 1991, using the Neftçi method, finds the timing of the probability signals from the Department of Commerce index of leading indicators to have been on average coincident with U.S. growth cycle turning points in 1951–86.[33] The individual signals show considerable dispersion, however, particularly at peaks. The 1962–64 growth cycle was missed. On the whole, the slowdowns and speedups might often have been recognized with relatively little delay but seldom detected early. Much the same can be said about the applications of the method to Japan and West Germany (where short lags prevail) and the United Kingdom (where the lags are longer).

In short, the probability approach has a broad appeal a priori on general statistical grounds and helps extend the range of uses of the leading indexes. Still, its results can only be as good as the predictive potential of these indexes allows. This is true as well for the method of sequential signaling by leading and coincident index growth rates, which retains its usefulness.[34]

11.3.4 Regression-Based Approaches

Leading indexes and their components can be used to predict not only turning points but also future values of macroeconomic aggregates with coincident timing tendencies. For example, consider the ordinary regression of dq_t on $d\ell_{t-1}$, where d denotes first differences in logs ($\Delta \ln$) and q refers to real GNP and ℓ refers to the index of leading indicators. Let the parameter estimates based on quarterly data for 1948–69 be used to compute one-quarter-ahead forecasts of real growth in 1970–76, and then reestimate for 1948–70 and 1971–76 and so on. As shown in Vaccara and Zarnowitz 1978, this very simple procedure yields root mean square errors of about .01 (close to the standard errors of regression) and mean errors near 0. The correlations between the out-of-sample predictions and realizations range from .68 to .74. Adding dq_{t-1}, $d\ell_{t-2}$, and $d\ell_{t-3}$ does not improve the results significantly, but adding lagged changes in the inverted lagging index, dv, does. A chain of predictions can extend the span of forecasts of real growth: dq_{t+2} is obtained from $d\ell_{t+1}$, which is itself obtained from dv_t. Measures of relative accuracy

33. Whether the confidence levels of 90% or 95% are used apparently makes little difference, and Niemira opts for the latter. The growth cycle chronology is that used in studies of the NBER and the Center for International Business Cycle Research (CIBCR). Comparisons with other chronologies such as Beveridge and Nelson 1981 are shown to yield much worse results.

34. See section 11.3.2 above on the application of this method to U.S. business and growth cycles. Reports by Moore on the continuing work at the CIBCR indicate that applications to other countries also provide useful information on their growth cycle developments.

suggest that the forecasts with composite indexes are definitely "in the ball park," that is, not at all far behind the much more complex and expensive GNP forecasts of econometricians and business economists.[35]

Auerbach (1982) estimates jointly (1) the regression of a cyclical variable *dy* on its own lagged values and on those of the leading index, $d\ell$, and (2) the regression of $d\ell$ on its component series. Monthly rates of change in either the unemployment rate or the industrial production index serve as *dy*, monthly rates of change in the BEA index of 1975 vintage serve as $d\ell$, and the lags range from 1 to 10 months. The weights estimated from Auerbach's equation (2) for the sample period 1949–77 vary considerably across the index components and differ depending on which of the two cyclical variables is used. The regression weights help to obtain a better within-sample fit for the unemployment rate, but the BEA index with approximately equal weights performs better in out-of-sample predictions for both variables. These results prove robust to different choices of sample period (1949–73, 1968–73) and forecast period (1973–77 and subdivisions). The stability of equations is rejected when using the estimated index with "optimal" regression weights but accepted when using the BEA index. Exclusion of those component series that did not individually help explain *dy* worsens the out-of-sample prediction performance of the BEA index.

Leading indicators contribute to the determination of national output and some other major economic aggregates in vector autoregressive (VAR) models, as shown in chapter 12. Continuing work in this area, conducted by me jointly with Phillip Braun, is now producing updated regression estimates and out-of-sample predictions with selected indexes. The latter include combinations of new-investment commitments for plant and equipment and housing (ℓ_2) and of most of the other components of the BEA leading index used in the late 1970s and 1980s (ℓ_8). Table 11.8 compares forecasts from four VAR models. The first model (A) includes real growth, inflation, money growth, and the interest rate; another (B) adds ℓ_2; the third (C) adds ℓ_8 instead; and the fourth (D) includes all six variables. The sample period is 1949–82; the forecast period is 1983–89.

The ratios of model B to model A root mean square errors (RMSE) show that when ℓ_2 is included, forecasts of real growth (q) improve for the first three quarters but deteriorate for longer spans. Interestingly, ℓ_2 helps to raise the accuracy of forecasts of inflation (p) and the level of interest rates (I) at all horizons (cols. 1–3). The RMSE ratios for C/A tell us that ℓ_8 contributes to the forecasts of q over quarters 1–4, to those of p over some intermediate and longer spans, and to those of I over most spans. However, most of the C/A

35. Leading indicator predictions of changes in the coincident index (*dc*) were on the whole more accurate than the predictions of *dq*, according to the results by Vaccara and Zarnowitz. One possible reason is that the coincident index has smaller measurement errors than real GNP; another, that it has more of a tendency to lag slightly. (Errors in component series often average out in an index, and employment lags slightly behind output much of the time.)

Table 11.8 Measures of Net Predictive Value of VAR Models with and without Selected Leading Indexes: Forecasts for 1–12 Quarters Ahead, 1983–89

	Ratios of Root Mean Square Errors (%)								
	Model B/Model A			Model C/Model A			Model D/Model A		
Span in Quarters	*q* (1)	*p* (2)	*I* (3)	*q* (4)	*p* (5)	*I* (6)	*q* (7)	*p* (8)	*I* (9)
One	81	83	93	93	100	92	83	80	91
Two	86	82	96	96	103	100	96	81	102
Three	76	89	96	95	99	97	82	89	101
Four	101	90	92	98	98	94	99	90	92
Five	114	92	92	110	100	81	117	90	89
Six	125	94	94	110	99	94	122	91	93
Nine	102	94	95	123	97	94	117	92	91
Twelve	102	96	97	107	99	95	101	96	92

Note: Model A includes four variables: q, m, I, and p. Model B includes the same variables plus l_2. Model C replaces l_2 with l_8. Model D includes all six variables: q, m, I, p, l_2, and l_8. The variables are defined as follows ($\Delta\ln$ = first difference in natural logarithm): $q = \Delta\ln$ real GNP; $m = \Delta\ln$ money supply M1; I = commercial paper rate; $p = \Delta\ln$ GNP implicit price deflator; $l_2 = \Delta\ln$ index of new-investment commitments (contracts and orders for plant and equipment; building permits for private housing); $l_8 = \Delta\ln$ index of eight leading indicators (average workweek, manufacturing; average initial claims for state unemployment insurance; vendor performance—percentage of companies receiving slower deliveries; change in sensitive materials prices; new orders for consumer goods and materials, constant dollars, manufacturing industries; index of net business formation; change in inventories on hand and on order, constant dollars, manufacturing and trade; index of 500 common stock prices). Cols. 1, 4, 7 refer to forecasts of real GNP growth; cols. 2, 5, and 8, to forecasts of inflation; cols. 3, 6, and 9, to forecasts of the interest rate.

ratios are higher than the corresponding B/A ratios, which suggests that the net predictive value of ℓ_8 is less than that of ℓ_2 (cf. cols. 1–3 and 4–6). When both ℓ_2 and ℓ_8 are included, some gains are achieved, particularly in forecasting p, as indicated by the D/A ratios (cols. 7–9). It might appear that the contributions of ℓ_2 and ℓ_8 are small (most of the ratios, in percentages, are in the 90s), but one should remember that the benchmark VAR model A is not an easy one to beat (consisting as it does of four principal macroeconomic variables, each quarterly series being taken with four lags, plus constant terms and time trends).[36]

Finally, Diebold and Rudebusch in their most recent evaluation (1990b) look at the contribution of the Department of Commerce leading index (LI) to monthly forecasts of industrial production (IP) that also use the lagged values of IP. They report that LI has a considerable net predictive value when both series (used in logarithms of the levels) are taken in the "final" revised form, but not when LI is the preliminary series and IP is the revised one. In the latter

36. The autoregressive elements are particularly strong for p, m, and I but relatively weak for q, l_2, and l_8, as indicated by F-statistics. The most significant influences on q are those of the interest rate and the leading indexes. Compare the similar results of the F- and t-tests reported in chapter 12.

case, of course, the past LI values would be available to the "real-time" forecaster but the past IP values would not. Here again, my interpretation of their results is that they primarily reflect the measurement errors in the preliminary LI series (as in Diebold and Rudebusch 1989; see sec. 11.3.3 above).

11.4 Recent and Expected Innovations

11.4.1 New Methods and Experimental Indexes

An experimental index of coincident indicators by Stock and Watson (1989, 1991a) is built on the assumption that all comovements of the included time series are attributable to a single unobserved factor (the "aggregate economic activity" or "state of the economy").[37] The underlying linear stochastic model is specified in log differences, so the common dynamic factor is $\Delta \ln C_t$; in addition, the relative change in each of the coincident variables covered is supposed to include a random component u_t that collects the variable-specific movements and measurement errors. The u's are thought of as uncorrelated with $\Delta \ln C_t$, with each other across the component series, and over time, at all leads and lags.

The Stock-Watson (S-W) coincident index includes four series, three of which it shares with the BEA index. One series is slightly different (employee-hours are used by S-W and number of employees by BEA, both for the nonfarm sector). The two indexes are very similar but the cyclical movements are somewhat smaller in the S-W composite and the upward trend is weaker (Stock and Watson 1989, p. 366).

The S-W formalization of the coincident index is methodologically interesting but rather restrictive. It certainly does not imply that business cycles have a single common cause, or that they can be reliably dated according to the movements of a single coincident index, BEA or S-W (cf. Stock and Watson 1989, pp. 352–55; Zarnowitz and Braun 1989, p. 399). A broader range of variables, nominal as well as real, would require more than one factor.

The new S-W leading index is both conceptually and structurally very different from the older NBER and BEA indexes. Calculated from a modified VAR model with seven selected leading variables, it is a forecast of the 6-month-ahead growth rate of the S-W coincident index ($\ln C_{t+6|t} - \ln C_{t|t}$). The selection involved an extensive search through a large array of series limited to the sample period 1960–88, using models with up to nine lags in each of the leading series and four lags in $\Delta \ln C$. Stock and Watson recognize that the danger of overfitting is particularly large here (1989, p. 367). A good in-sample fit would be expected and is largely confirmed (Zarnowitz and Braun

37. See Sargent and Sims 1977 for an early study of an unobserved one-index model of multivariate time series.

1989). But in the absence of out-of-sample tests, doubts arise about just how well this index would have performed over past periods with very different characteristics or how well it will do in the future.

The S-W leading index includes no series on money and credit, inventory investment, consumption, delivery lags, sensitive commodity prices, and stock prices. Three of its seven components relate to interest rates: change in the 10-year Treasury bond yield (I_{10}); the 6-month risk spread, commercial paper rate minus Treasury bill rate (CP – TB); and the yield curve, 10-year to 1-year Treasury bond spread ($G10 - G1$). The other components are the building permits for private housing units (BP); manufacturers' unfilled orders for durable goods, in constant dollars (DUO); the trade-weighted nominal exchange rate (EXN); and part-time work in nonagricultural industries due to slack (PTW). BP is a level series; the others are smoothed growth rates.

Only two of the seven S-W leaders are also included in the BEA index (BP and DUO, but the latter in the form of absolute change, not rate of growth). A series such as EXN, with a short and poor record of cyclical conformity and timing, would not have qualified under the selection criteria used in the indexes constructed in the past (Moore and Shiskin 1967; Zarnowitz and Boschan 1975a, 1975b). Although theoretical considerations support inclusion of this important variable, they do so equally for other variables that have been omitted, particularly those relating to real investment, profits, credit, and money.

The level of interest rates tends to move procyclically with lags (the prime rate is a component of the BEA index of lagging indicators), but as shown earlier, such laggers when inverted become long leaders. The I_{10} component contributes only mildly to the S-W index.[38] The yield curve ($G10 - G1$) is more effective, and the public-private spread (CP – TB) is by far the most effective, according to the tests reported by S-W and Zarnowitz and Braun 1989 (but see sec. 11.4.3 below for some qualifications).

The S-W leading index would probably gain from broader and more diversified coverage, including at least some of the more effective components of the BEA index. Its reliance on interest rates and spreads seems too heavy: the predictive power of these variables in the past was not always as strong as in recent times, and it may not always prove as strong in the future (Sims 1989, pp. 396–97). Indeed, the S-W index failed to predict the July 1990 business cycle peak. However, most leading indicators performed rather poorly on this occasion. (Stock and Watson 1991b), perhaps because the 1991 downturn was aggravated and accelerated (though not necessarily caused) by the Iraqi invasion of Kuwait on August 2. The BEA leading index declined in July, synchronously with the general economic activity.

38. See, e.g., fig. 5(F) in Stock and Watson 1989, p. 371. Also, note that "interestingly, including a measure of ex ante real rates (with various measures of expected inflation) does not improve the performance of the LEI [i.e., S-W] index" (p. 385).

11.4.2 Do Stock Prices Belong in the Leading Index?

Probably no major leading indicator met with more criticism in recent times than the stock price index (particularly after the crash of October 1987). Stock prices reflect profit expectations and presumably help determine consumption (wealth effect) and cost of capital, hence business investment decisions (*q*-ratio). Some studies offer strong support for the power of market changes to predict output (Fama 1981; Fischer and Merton 1984); others provide mixed evidence on the theoretical consequences of stock price movements (see references in chapter 2 and Cagan 1990). The verdict of most recent tests is essentially negative.[39]

I find the long-term cyclical record of stock prices as leading indicators in the United States and other developed economies (Canada, Japan, and Western Europe) to be good in the relevant relative terms. The market indexes have performed worse than the composite indexes, but this is true of other individual indicators as well. Their recent record leaves much to be desired, but it would seem premature to eliminate the S&P 500 index from the U.S. leading index. As summed up in table 11.9, all but three of the U.S. business contractions since 1873 were preceded and accompanied by substantial declines in the market.[40] So stock prices signaled well the actual recessions and recoveries. However, they performed poorly with respect to the other type of error by giving many false signals (i.e., containing many extra turns). Of the 12 significant declines in the S&P index in 1946–82, 5 (42%) were not associated with recessions. Interestingly, the conformity of stock prices to business cycles, as measured by both types of error, was excellent in 1873–1919 and much worse in the later era of the Federal Reserve and rising government size and interventions (cf. table 11.9, cols. 4 and 5). Yet one should also consider that the tendency of the market to lead actually strengthened in the postwar period (cols. 6–7). Also, the market declines associated with recessions and major slowdowns were typically larger and longer than the other declines.[41]

11.4.3 New Proposals and Need for Further Work

The study by Stock and Watson suggests that the BEA index could be significantly improved by the inclusion of some functions of interest rates, par-

39. Stock and Watson (1989, p. 388) conclude that the marginal predictive value of S&P 500 in helping to predict their coincident index is "modest." Fredman and Niemira (1990), working with the components of the latest BEA composite index of leading indicators, consider the same stock price series the main candidate for replacement. For a rather more optimistic appraisal, however, see Niemira 1990a.

40. The exceptions consist of the mild, indeed marginal, recession of 1926–27, the short one in the first half of 1980, and the 1945 episode related to the end of World War II.

41. The bear markets in 1962 and 1966 are related to business slowdowns; those in 1946, 1977, and 1983–84 are not. The outliers are the short but steep decline in 1962 and the long but gradual decline in 1976–78. The average amplitudes per month of cyclical contractions in common stock prices are close to 2% for each of the three subperiods distinguished in table 11.9. For more detail on timing and amplitudes of stock price indexes in the United States, see Moore 1983, ch. 9, and Zarnowitz 1987–88. On the timing of market indexes at postwar growth cycle turns in eight countries, see Klein and Moore 1985.

Table 11.9 **Selected Measures of Cyclical Conformity and Timing, Common Stock Prices, 1873–1986**

Period	No. of Turning Points Covered in Business Cycles (B.C.)[a] (1)	No. of Turning Points Covered in Stock Price Cycles (S.C.)[b] (2)	No. of B.C. and S.C. Turning Points Matched[c] (3)	% of B.C. Turns Skipped[d] (4)	% of Extra S.C. Turns[e] (5)	% of Leads[f] At B.C. Peaks (6)	% of Leads[f] At B.C. Troughs (7)
1873–1919	24	24	24	0	0	83	75
1920–1945	12	10	8	33	20	75	75
1946–1986	16	24	14	13	42	100	100
1873–1986	52	58	46	12	21	87	83

Sources: Morgenstern 1959, ch. 10; Moore 1983, ch. 9; *BCD*, April 1987, p. 105; Zarnowitz 1987–88, p. 14.

[a]As dated by the NBER. Every entry includes an equal number of B.C. peaks and B.C. troughs.

[b]Refers to specific-cycle peaks and troughs (in equal numbers for every entry) in the S&P monthly index of common stock prices. For 1873–1919: index compiled by Cowles Commission and S&P, includes virtually all industrial public utility, and railroad stocks traded on the New York Stock Exchange. For 1918–56: monthly average of S&P weekly index, coverage increasing from 198 to 480 stocks. 1957–86: S&P index of 500 common stocks.

[c]B.C. turns matched by like turns in the monthly S&P 500 index.

[d]Equals in terms of the column numbers [(1) − (2)/(1)]100.

[e]Equals in terms of the column numbers [(2) − (3)/(2)]100.

[f]Based on all timing observations at matched B.C. and S.C. turns (see cols. 3 for the numbers and nn. *a–c*).

ticularly the risk spread. Some measure of the slope of the yield curve is another candidate that receives support from both historical analyses (Kessel 1965; Cagan 1966) and recent tests (Keen 1989; Fredman and Niemira 1990).

Corporate bond prices led at each peak and each trough of the eight complete business cycles of 1948–82, but they have a much longer record of consistent cyclical behavior (Mitchell and Burns 1938). The Dow-Jones index, a simple average of daily closing bond prices for each month, had very long and highly variable leads at peaks (with a mean and a range of 27 and 10–58 months, respectively). Its leads at troughs were more moderate but also long relative to the observed distributions of such leads and the durations of business contractions (they ranged from 3 to 13 months and averaged 7 months). Bond prices, of course, move inversely to bond yields. Their long downward trend in the postwar period reflects the rise in the price level and the gradual diffusion of inflationary expectations.[42]

The Dow-Jones bond price average is included in the new long-leading index of the CIBCR along with the money supply (M2) in constant dollars, new housing permits, and the ratio of the price to unit labor cost in manufactur-

42. For detail on corporate bond prices as a leading indicator, see Zarnowitz 1990.

ing.[43] The long-leading index is designed to help remedy the old problem of very short leads at business cycle troughs of the existing composites. It is also intended to provide a forecasting tool with a somewhat longer forward reach and to be tested against the movements in the short-leading index (and vice versa). Its peaks (troughs) lead the corresponding turns in the BEA index on average by 4 (5) months (see Moore and Cullity 1990).

Seven of the 11 components of the new short-leading CIBCR index are the same as or nearly equivalent to the indicators included in the BEA leading index in its present form. The remaining components include a new index of net business formation, a new layoff rate series, growth rate in real domestic nonfinancial debt, and the diffusion index of inventory change from the monthly survey of the National Association of Purchasing Management. Each of these series represents an important cyclical variable with a tendency to lead well known from the results of closely related data used in the past.[44]

With larger samples of better data and a careful choice of evaluation methods, we should be able to draw important distinctions between types of fluctuations and turning points. As noted earlier, some indicators lead only at peaks, others only at troughs, and neither of these groups should be ignored. Also, some indicators may rank high as predictors of business cycle turns but low as predictors of growth cycle turns, or vice versa.[45]

Much more work needs to be done to extend and improve the set of useful indicators of major changes in general economic activity. This applies not only to the leading series but also to the roughly coincident and lagging series.

43. The first two of these series are also components of the BEA leading index (with the minor difference that BEA uses an index of permits, and CIBCR uses the number of permits). The ratio series is a new monthly indicator of profit margins (see Boschan and Moore 1990).

44. For analyses of these indicators, see Moore 1990, ch. 2; Moore and Cullity 1990; Klein and Moore 1990; and Zarnowitz and Moore 1990. For criticism of the two series included in the BEA index but not in the CIBCR indexes (change in constant-dollar unfilled orders for durable goods and the index of consumer expectations), see Moore 1990, pp. 12–14.

45. Fredman and Niemira 1990 evaluate 17 leading indicators separately at the two sets of turns in 1948–82, using a Neftçi-type technique. In signaling business cycle turns, the ratio of the coincident to the lagging index ranks first, the yield curve second, I_{10} fourth, the BEA leading index fifth, and the "credit risk" (CP − TB) last. In signaling growth cycle turns, the BEA leading index ranks first, credit risk fifth, the yield curve eleventh, I_{10} fifteenth, and the C/Lg ratio last!