SAUL H. HYMANS
University of Michigan

On the Use of Leading Indicators to Predict Cyclical Turning Points

THE POTENTIAL VALUE OF ACCURATE and timely information on current and developing economic conditions is huge. While the value may differ for a legislator, a corporate president, the chairman of the Federal Reserve Board, a Treasury Department revenue estimator, or an econometric forecaster, such information carries extremely high value—both private and social. Official acceptance of this fact dates back at least to 1957, when Raymond J. Saulnier, then chairman of the Council of Economic Advisers, made a solicitation to the Bureau of the Census that resulted in the publication in 1961 of the monthly *Business Cycle Developments*.[1] As first published, *BCD* took maximum advantage of the long years of painstaking research by Wesley Mitchell, Arthur F. Burns, Geoffrey H. Moore, Julius Shiskin, and others—many of them associated with the National Bureau of

Note: The research underlying this paper was supported in part by National Science Foundation Grant GS-36932X. The author wishes to express his gratitude to Robert S. Holbrook, Ronald D. Lee, Stephen Marston, Harold T. Shapiro, Lester D. Taylor, and several members of the Brookings panel for their valuable aid at various stages of the research. The author is especially grateful to Jesse Hall, who provided substantial help throughout all aspects of the project and who was, in short, the model of a dedicated research assistant.

1. The background of this event and the technical rationale underlying the data published in *BCD* are discussed in Julius Shiskin, *Signals of Recession and Recovery: An Experiment with Monthly Reporting*, Occasional Paper 77 (National Bureau of Economic Research, 1961). In November 1968 the document was renamed *Business Conditions Digest*, but retained its *BCD* initials.

Economic Research—who favored the so-called "cyclical indicators approach" to the analysis of business conditions. The major feature of the early volumes of *BCD* was the regular publication of some eighty economic series that had been judged to be important measures or indicators of national economic conditions. More than fifty of these series had been classified into "leading," "roughly coincident," or "lagging" groups, according to their typical timing and conformity with respect to U.S. business cycles.[2]

During the past half dozen years, the cyclical indicators data have lost their monopoly position in *BCD*, which has come to include a vast array of income and product account data, anticipations and intentions data, diffusion indexes, and other measures. This shift reflects the fact that serious students of economic conditions use a variety of analytical approaches to the subject, including the econometric model, and that a more inclusive data base would consequently be better able "to provide statistical information so arranged as to facilitate the analysis of the course of the Nation's economy" for a broader group of users.[3] At the same time, the National Bureau of Economic Research (NBER) itself has recently come to the view that the analysis and forecasting of economic conditions can profitably be undertaken in a variety of formats, with varying emphases and advantages. Under NBER auspices, Victor Zarnowitz, Jacob Mincer, Rosanne Cole, and others have done seminal work on the evaluation of forecasting accuracy and the usefulness of various kinds of forecasting techniques based on autoregressive time series analysis as well as standard macroeconometric models.[4] Although these developments have to be read as a *relative* diminution of interest in the cyclical indicators approach, no evidence suggests that the indicators are about to disappear. Indeed, the 1960 list of indicators that formed the original basis for *BCD* was extensively restudied under the direction of Moore and Shiskin, and a revised and expanded list

2. See Shiskin, *Signals of Recession*, pp. 15–17.

3. See the statement to this effect on the inside front cover of any recent issue of *BCD*.

4. See, for example, Victor Zarnowitz, *An Appraisal of Short-Term Economic Forecasts*, Occasional Paper 104 (Columbia University Press for the National Bureau of Economic Research, 1967); and Jacob Mincer (ed.), *Economic Forecasts and Expectations: Analyses of Forecasting Behavior and Performance*, Studies in Business Cycles 19 (Columbia University Press for the National Bureau of Economic Research, 1969).

—the 1966 list—was published in 1967.[5] Shiskin and others have continued their efforts to improve the currency and predictive content of the indicators and, to this end, have developed and refined a number of techniques for removing the so-called "noise component" from a time series.

In recent years, judging by speeches and press releases, government officials have become increasingly convinced of the predictive reliability of the indicators. Probably at least some individuals with responsibility for stabilization policy treat the leading indicators as a major input to their own recommendations. The media have obviously sensed this growing interest and have themselves latched on to the cyclical indicators. The "short list" of leading indicators, released monthly, is now accorded publicity perhaps greater than its significance warrants, and surely beyond the consuming public's ability to place it in proper perspective.[6]

The very facts of official publication and growing interest by government and the press impart credibility and respectability to the cyclical indicators. But what of the cold, objective evidence? Does it point to the indicators as a reliable foundation for important public and private decisions? There are two distinct bodies of literature on which to base a scientific conclusion about the value of the indicators. The first is a large and growing literature, produced by the proponents of the indicators themselves, which represents a continual effort to monitor their performance.[7] These studies accept the validity of the indicator approach to forecasting and concentrate on interpreting recent movements in the indicators, improving measurement, and updating the indicator lists. An entirely separate body of evidence consists of studies by disinterested parties. Many, but not all, of these express frank skepticism about the indicator approach and succeed in turning up

5. Geoffrey H. Moore and Julius Shiskin, *Indicators of Business Expansions and Contractions*, Occasional Paper 103 (National Bureau of Economic Research, 1967).

6. This situation is in no way unique to the leading indicators. It occurs when the preliminary GNP accounts are published quarterly; in the November–December period when many econometric model forecasts of the coming year come out in rapid-fire succession; when the quarterly reports on consumer sentiment are issued by the Survey Research Center of the University of Michigan; and so on.

7. The most recent additions in this area are two papers by Geoffrey H. Moore: "New Work on Business Cycles," prepared for the *53rd Annual Report* (1973) of the National Bureau of Economic Research (forthcoming); and "Economic Indicator Analysis During 1969–72," to appear in *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz* (forthcoming).

rather damaging evidence. Unfortunately, most of the best of these studies are more than a decade old. Indeed, I know of only one new study that belongs in this group and that has evaluated any of the recent data in depth.[8]

I submit, then, that the degree of interest in the indicators has outstripped the production of objective, "outside," evidence bearing on their scientific merit. Such a situation seems to me to call for a new appraisal. A general reappraisal constitutes a research project worthy of at least one large volume and clearly well beyond the scope of a single paper. I have therefore been forced to define a narrower study which could lay some claim to having been conducted in depth without severely stretching the bounds of a single, readable paper. In such a situation priorities become extremely important, and I believe that an evaluation of the construction and forecasting content of the "composite index of leading indicators" represents the first order of business. Even within this narrow focus, some may be disappointed by an absence of comparisons between the forecasting accuracy of the leading indicators and existing macroeconometric models, and by the relative emphasis that I place on qualitative, rather than quantitative, forecasting accuracy.[9]

8. This study is H. O. Stekler and Martin Schepsman, "Forecasting with an Index of Leading Series," *Journal of the American Statistical Association*, Vol. 68 (June 1973), pp. 291–96. I was unaware of their paper until after I had written the original draft of this paper and have since read their work with great interest. Their principal finding anticipates one of the results of my section on turning point tests, but beyond that the two papers are quite distinct. My own work shows the influence of the 1962 analysis done by M. Hatanaka, "Application of Cross-Spectral Analysis and Complex Demodulation: Business Cycle Indicators," which appears as Chapter 12 in C. W. J. Granger, in association with M. Hatanaka, *Spectral Analysis of Economic Time Series* (Princeton University Press, 1964). Well-known earlier studies were done by Arthur M. Okun, "On the Appraisal of Cyclical Turning-Point Predictors," *Journal of Business*, Vol. 33 (April 1960), pp. 101–20; and Sidney S. Alexander, "Rate of Change Approaches to Forecasting-Diffusion Indexes and First Differences," *Economic Journal*, Vol. 68 (June 1958), pp. 288–301. A useful reference for the literature in the area is D. J. Daly, "Forecasting with Statistical Indicators," in Bert G. Hickman (ed.), *Econometric Models of Cyclical Behavior*, Studies in Income and Wealth 36 (Columbia University Press for the National Bureau of Economic Research, 1972), Vol. 2.

9. With regard to a comparison with macroeconometric models, the major problem is that there is not yet a sufficient history of quarterly econometric forecasts by a sufficient number of models to render such a comparison especially useful. On the matter of qualitative vs. quantitative forecasting accuracy, my choice reflects a personal belief that forecasting of directions of change or of turning points should be regarded as the avowed purpose of the indicators, and that success in that sphere alone would constitute a major contribution.

The next section of the paper identifies the twelve variables that form the composite index of leading indicators (hereafter CLI) and presents my view of the rationale for constructing the CLI. The following section contains the results of a number of turning point tests conducted on the CLI, and the next applies the technique of spectral analysis to evaluate the manner of construction of the CLI and to shed additional light on the results of the turning point tests. There follows an indication of how spectral analysis might be used to form a new leading indicator that overcomes some of the difficulties revealed in the current CLI, and two variants of such an indicator are then evaluated. Conclusions and suggestions for further work are contained in the final section.

The Composite Leading Indicator and Its Components

The current CLI is constructed in three basic steps which produce a single monthly index from the 1966 short list of twelve leading indicators.[10] The twelve components are listed with abbreviated titles in Table 1, and are thereafter referred to in the text and subsequent tables by even briefer titles. Column 2 in the table records the median lead of each series with respect to the turning points in the NBER reference cycles from the year shown in column 1 through 1965, the terminal point for construction of the 1966 short list. Column 4 updates column 2 with the median lead at reference turns in the 1948–70 period; and columns 5 and 6 display the range of leads and lags at peaks and troughs separately for the more recent period. Column 4 displays substantially more uniformity in the median lead than is shown in column 2. Columns 5 and 6, however, reveal that the timing of each leading indicator displays a very wide range of variation at reference turns. This matter will receive attention in subsequent discussions; suffice it to say here that this variability causes obvious difficulties in using the specific indicator variables or the CLI itself to predict turning points.

The first step in constructing the CLI is to standardize the monthly data so that each series displays the same average absolute monthly change.[11] This makes it possible to combine series with inherently different units of measurement. The next step is to combine the adjusted series into a weighted

10. One short list variable, corporate profits after taxes, is actually a quarterly series. The details of its incorporation into the monthly CLI need not concern us here.

11. I shall attempt only to provide the flavor of the main points in the construction of the CLI. A fuller description can be found in Moore and Shiskin, *Indicators*.

Table 1. Selected Characteristics of the 1966 Short List of Leading Indicators Used to Construct the Composite Leading Indicator, Subperiods through 1970

BCD series	*Period through 1965*			*1948–70*		
	First cyclical turn covered (*1*)	*Median lead* (−) *or lag* (+) (*months*) (*2*)	*Index score* (*3*)	*Median lead* (−) *or lag* (+) (*months*) (*4*)	*Range of leads* (−) *or lags* (+) *Peaks* (*5*)	*Troughs* (*6*)
Average workweek, manufacturing	1921	−5	66	−5	−4 to −20	0 to −6
Average weekly initial claims, unemployment insurance (inverted)	1945	−6	73	−5.5	−10 to −22	+1 to −1
New orders, durable goods	1920	−4	78	−3.5	−2 to −13	−1 to −5
Contracts and orders, plant and equipment	1948	−6	64	−5.5	−5 to −10	+3 to −6
Net business formation	1945	−7	68	−7	−9 to −25	0 to −5
Corporate profits after taxes	1920	−2	68	−5.5	−2 to −20	0 to −9
Ratio of price to unit labor cost, manufacturing	1919	−3	69	−5	−5 to −30	0 to −8
500 common stock prices	1873	−4	81	−5.5	−5 to −12	−4 to −11
Industrial materials prices	1919	−2	67	−5	+3 to −29	+8 to −10
New building permits	1918	−6	67	−9.5	−8 to −29	−2 to −11
Change in book value, manufacturing and trade inventories	1945	−8	65	−5.5[a]	−4 to −15	0 to −9
Change in consumer installment debt	1929	−10	63	−6.5	−6 to −28	+2 to −9

Sources: Cols. 1–3—Geoffrey H. Moore and Julius Shiskin, *Indicators of Business Expansions and Contractions*, Occasional Paper 103 (National Bureau of Economic Research, 1967), Table 6, pp. 36–45; cols. 4–6—U.S. Bureau of Economic Analysis, *Business Conditions Digest*, Series ES1, No. 73-3 (March 1973), Table F, p. 115. Computations of medians by author.

a. Applies to the period 1948–61, there having been no discernible specific cycle corresponding to the peak and trough in 1969 and 1970.

average using the index scores shown in column 3 of Table 1. Thus, the index of common stock prices receives the largest weight in the CLI, while the change in installment debt receives the smallest. These scores, or weights, represent the average score received by each series on six criteria, each of which has a potential score from zero through 100. The six criteria test such things as the "economic significance" of the series, its "conformity" and "timing" with respect to the reference cycle, how quickly it becomes available (its "currency"), and so on. The resulting average is then subjected to a "reverse trend adjustment" to ensure that the index has precisely the same trend as the composite index of coincident indicators (hereafter, CCI).[12] The CCI, like the gross national product or industrial production, has a pronounced upward trend. A number of the series in Table 1 relate more to output changes than to output levels and have little if any trend. The first two steps in the process of constructing the CLI therefore produce a series that displays many small declines *not* indicative of a coming drop in the CCI. The reverse trend adjustment is intended to overcome this difficulty.

At this point an obvious question suggests itself: Having identified the series to be combined into the CLI, why not run a regression and let the data themselves decide on the weights to be applied?[13] To do this, of course, requires the designation of a dependent variable to be regressed on the independent variables contained in Table 1. Since the purpose of the leading indicators is to anticipate the movements in general business conditions, either the CCI itself or its price-deflated version are obvious candidates for the dependent variable.[14] The results of a multiple regression explaining the

12. See Julius Shiskin, "Reverse Trend Adjustment of Leading Indicators," *Review of Economics and Statistics*, Vol. 49 (February 1967), pp. 45–49.

13. In fact, John E. Maher posed precisely this question and then used regression analysis to construct an index in "Forecasting Industrial Production," *Journal of Political Economy*, Vol. 65 (April 1957), pp. 158–65.

14. Solomon Fabricant has recently argued that cyclical reversals in the direction of economic activity ought to be defined in terms of "real" or price-deflated variables rather than a mixture of data in constant and current prices, as has been the tradition. My basic agreement with this position has led me to use the deflated CCI as the standard to be predicted or explained throughout this study. I have, however, made no attempt to deflate the components of the CLI. I am not certain, in many cases, what deflator ought to be used, nor am I certain that the argument that applies to the variables that *define* the cyclical turns necessarily applies to all variables that are to be considered *indicators* of the cyclical turns. See Solomon Fabricant, "The 'Recession' of 1969–1970," in Victor Zarnowitz (ed.), *The Business Cycle Today* (Columbia University Press for the National Bureau of Economic Research, 1972).

Table 2. Regression of the Deflated Composite Coincident Indicator on the Short List of Leading Indicators[a]

Variable and median lead (months)	*Regression coefficient*	*t-statistic*
Workweek (−5)	−1.479	−3.49
Unemployment claims (−6)	2.864	8.20
Durable orders (−4)	1.461	10.56
Plant and equipment contracts (−6)	0.902	2.10
Business formation (−7)	−0.001	−0.04
Profits (−2)	0.294	5.24
Price-cost ratio (−3)	−0.199	−1.80
Stock prices (−4)	−0.027	−0.98
Materials prices (−2)	−0.043	−2.53
Building permits (−6)	−0.019	−2.41
Inventory change (−8)	−0.129	−4.89
Installment debt (−10)	−0.026	−0.46
$\bar{R}^2 = 0.991$ Standard error of estimate = 2.59		

Source: Derived by author from Table 1.
a. The regression included an intercept and a quadratic trend and was fitted to monthly data over the period from 1948:1 through 1972:2.

deflated CCI by the variables in Table 1, each lagged by the number of months shown in column 2 of Table 1, is reported in Table 2.[15] One may compare these results with a regression of the deflated CCI on the CLI, the latter lagged five months to correspond to its median lead at turning points. The resulting equation is

$$\begin{aligned} (1) \quad CCI = {} & \underset{(4.50)}{8.140} + \underset{(5.95)}{0.052}\, Time - \underset{(-7.26)}{0.004}\, (Time)^2 \\ & + \underset{(19.31)}{0.986}\, CLI_{-5}. \end{aligned}$$

$\bar{R}^2 = 0.984$; standard error of estimate = 2.85.

The numbers in parentheses here and in subsequent equations are *t*-statistics. It is clear that the relative weights that translate the short list variables into the CLI do not correspond to those chosen by a multiple regression of the CCI on the entire set of variables. In fact, according to the results in Table 2, three of the identified leading indicators appear to make

15. The results are qualitatively similar whether the CCI, the deflated CCI, or components of the CCI, such as industrial production or manufacturing and trade sales, is the dependent variable. The same is also true of minor variations in the lag pattern of the leading indicators.

no significant marginal contribution to the explanation of the CCI, and of the nine significant indicators, fully five appear with negative signs in the multiple regression equation![16]

These findings point up the conceptual differences between the CLI type construction and the predicted-value index yielded by a multiple regression equation. From the point of view of the latter, a prediction error of a given size carries the same weight no matter at what sample point it occurs. The regression of the CCI on the leading indicators does not represent a behavioral or technological relation; it is merely an exercise in curve fitting or—at best—some kind of pseudo reduced-form equation. The least squares procedure will assign whatever weights (coefficients) are required to fit the CCI with minimum residual variance over the observed sample of points. A poor fit at the turning points can be overwhelmed by a sufficiently close fit at the vast majority of data points between the turns in economic activity. If there were no essential distinctions between the empirical relations among the variables in turning point zones and in periods of fairly smooth unidirectional change, one would expect a multiple regression approach to be optimal. But the very point that has been argued most strongly by the proponents of the indicator approach—going all the way back to the early work of Mitchell—is that there *is* something special about the turning point zones. The CLI is alleged to be constructed so as to maximize the use of the turning point information contained in the various leading indicator series.[17] The CLI construction must be interpreted along the following lines:

1. A turn in overall business activity is generally preceded by a turning point in many of a group of series that can be identified.[18]

16. A similar sign phenomenon was found by Maher in "Forecasting Industrial Production."

17. Whether the CLI—*as constructed*—does indeed extract the maximum amount of turning point information from its component series is a matter that will be taken up below, but that is clearly the *intent*.

18. This statement does *not* imply direct causality (if it did, one would attempt to estimate a behavioral or technological relation that could be expected to hold outside the sample, that would have directly interpretable coefficients, and so on). Rather, it implies something about the process through which those forces that *do* lead to turning points operate within the structure of the U.S. economy. Thus a downturn that might result from a tightening of monetary policy would inevitably have an impact on some or all of the variables durable orders, plant and equipment contracts, profits, price-cost ratio, stock prices, building permits, and installment debt before it affected the aggregate level of production.

2. Many of the leading series will therefore signal any impending turn in overall activity.

3. Which of the series signals earliest and most strongly depends, however, on the real cause of the impending turn and the exact process through which it operates to induce the turn.

4. It is therefore necessary to provide a mechanism that gathers many potentially duplicative signals of the same impending event—that is, to provide for the many possible causes of a turning point, to judge the quantitative importance of events by "counting up" the number of potential indicators giving the same signal, and thereby to reduce the likelihood of the index being overly affected by a false signal.

5. This implies not the fitting of a regression plane, but the averaging of a broadly based group of conforming series with positive weights somehow representative of their individual reliability in signaling turning points.[19]

While I am not entirely certain that the proponents of the CLI would subscribe to every detail of the preceding paragraph and its five summary points, I do believe that it represents a fair inference of their view and that it is a defensible position.[20]

Turning Point Tests of the CLI

The CLI exists primarily to signal the direction of change of overall economic activity, and its success must be judged largely in that context. Unfortunately, there exists no "User's Guide to the CLI" and therefore no prescribed methodology by which the message of the indicator is supposed to be read. The officially published leads and lags of the CLI at specific reference turns are all established after the fact, by hindsight, on the basis of the most recent data. Table 3 contains the record of the CLI, on this

19. This is a rather unusual "model" in which combining a number of simple regressions may indeed be preferable to running a multiple regression. A weighted multiple regression—that is, one that weighted the turns more heavily than the rest of the data points—might accomplish the same thing. But considerable artistry would be required to devise an appropriate set of weights.

20. Again, I mean this to be interpreted as defensible in concept; I am not implying that the CLI cannot be improved upon, even in the context stated above. It might also be remarked here that the CLI did prove to be generally superior to the predicted values from the equation in Table 2 in the kind of turning point tests to be reported on in the next section.

Table 3. Official Leads and Lags of the Composite Leading Indicator at Cyclical Turning Points, 1948–70

Peak	*Lead (−) or lag (+) (months)*	*Trough*	*Lead (−) or lag (+) (months)*
November 1948	−5	October 1949	−5
July 1953	−6	August 1954	−9
July 1957	−8	April 1958	−2
May 1960	−4	February 1961	−2
November 1969	−2	November 1970	−1

Source: *Business Conditions Digest* (March 1973), Table F, p. 115.

ex post basis, for each of the officially recognized turning points since 1948. The record appears to be quite good at first glance, but a careful look at the CLI time series itself reveals several problems. Consider, for example, the following string of consecutive monthly changes in the CLI from July 1956 through August 1957:

Year and month	*Change in the CLI*	
1956 July	−0.4	
August	+1.1	
September	−0.2	
October	+0.3	
November	+0.7	←CLI official peak month
December	−0.1	
1957 January	−0.3	
February	+0.1	
March	+0.1	
April	−0.6	
May	+0.3	
June	+0.2	
July	−0.1	←Reference peak month
August	+0.2	

Is it really so obvious that the CLI reached its peak in November 1956 and not, say, March 1957 or June 1957? Once it is known that the reference peak occurred in July 1957, it becomes much easier to interpret the CLI as having been on a "downward wriggle" from November 1956 on, but until hindsight was available, such an interpretation would have been far more difficult both to arrive at and to sustain.

The point is that a test of the CLI must be based on a specific set of rules devised to determine, ex ante, the meaning of the signal being given. The rules must (a) provide a way to filter out false signals; (b) not filter so thoroughly as to reduce measurably the lead time of genuine signals; and (c) be simple enough to be operational. After some thought and experimentation, I have settled on the following set of rules:

1. Forecasts are made one month at a time, the possible forecasts being *no change* in the cyclical phase (NC), a *peak* is coming (P), or a *trough* is coming (T).

2. The signal obtained as of month t determines the forecast (either NC, P, or T) for month $t + 1$.

3. During an upswing phase, two consecutive declines in the CLI lead to the prediction P for the month following that in which the second consecutive decline occurs.

4. Once two consecutive declines of the CLI have been observed in an upswing phase, two consecutive increases in the CLI are required to define a false peak signal and change the prediction from P to NC.[21]

5. The rules during a downswing phase are precisely symmetric, with two consecutive increases in the CLI being required for a T prediction; and following a T prediction, two consecutive declines are required to define a false trough and change the prediction from T to NC.

Notice the requirement in the rules that the forecaster know whether the economy is currently in an upswing or downswing phase. For the purpose of the tests to be conducted, I shall assume that the NBER identifies and announces turning points with perfect accuracy and minimum possible delay. This means, for example, that if a peak is to be dated as having occurred at month t, the NBER proclaims the peak in month $t + 2$, just before the forecaster has to interpret the CLI signal of month $t + 2$.[22]

What remains now is to specify a set of requirements by which the performance of the CLI may be judged. I have settled on the following:

1. If a peak occurs in month t, the correct and required predictions for months $t - 3$, $t - 2$, $t - 1$, t, $t + 1$, and $t + 2$ are all P. That is, a three-

21. This means that two consecutive declines in the CLI will always be followed by at least two months of peak predictions.

22. For example, the NBER observes the CCI for month t in month $t + 1$ and finds that it increased. In month $t + 2$, it observes the CCI for month $t + 1$ and finds that it declined. Month t is immediately proclaimed to have been the peak month and the forecaster knows this as he reads the CLI of month $t + 2$.

month lead at the peak is *required* and, in view of the assumption made about identifying a peak, predictions of P must also be required for the two months following the peak.

2. Since it makes little sense to penalize a lead that is longer than three months, a P prediction is also considered to be correct as early as nine months prior to the peak month *provided* that it is part of a consecutive string of P predictions continuing at least through the peak month.[23]

3. The trough rules are similar except that the *required* lead is two months rather than three, and the maximum permissible lead is six months rather than nine. Because the official downswing phases since 1948 have all been about a year in length, it seems unreasonable to regard a lead longer than six months as informative.

4. For all months not covered under rules (1)–(3), the correct prediction is NC.

Before applying these procedures to test the CLI, a few more ad hoc decisions require explanation. The mini-recession of 1967 is not officially recognized, but it did contain a quarter in which real GNP fell and it is picked up by the CLI. I shall therefore count the mini-recession in the turning point tests, and for this purpose I have dated the peak and trough months as December 1966 and March 1967, respectively.[24] The 1959 steel strike and the 1970 auto strike show up very clearly as strong negative changes in the CLI. The former would lead to a false peak prediction; the latter would reverse the appropriate recognition of the 1970 trough. I have simply not permitted these false predictions to take place, on the grounds that anyone using the CLI at the time would (or at least should) have been able to discount these events.

Table 4 contains the record of the CLI at the turning points during the

23. Thus, no credit is given for an isolated (perhaps random) P prediction six months before the peak; to be counted as correct it must be followed consecutively by at least six more P predictions. If the peak month is month *t*, the following string of predictions is considered to have provided a one-month lead at the peak:

Month	$t-7$	$t-6$	$t-5$	$t-4$	$t-3$	$t-2$	$t-1$	t
Prediction	NC	P	P	NC	NC	NC	P	P

The earlier P predictions followed by the NC predictions are regarded as having been contradictory and misleading.

24. My dating of the turns is based loosely on the monthly behavior of the CCI and its components, and other evidence such as the Business Week index. In view of the shortness of the downswing phase in this case, I could require of the CLI nothing stronger than coincidence at the trough.

Table 4. Unofficial Leads and Lags of the Composite Leading Indicator at Cyclical Turning Points, 1948–70

Peak	*Lead (−) or lag (+) (months)*	*Trough*	*Lead (−) or lag (+) (months)*
November 1948	−8	October 1949	−2
July 1953	−3	August 1954	−6
July 1957	missed	April 1958	+1
May 1960	−1	February 1961	+1
December 1966	−6	March 1967	+2
November 1969	+1	November 1970	−3

Source: Calculations by author.

1948–70 period based on the rules and requirements specified above. A comparison of Tables 3 and 4 reveals some noticeable differences. The official data record an ex post eight-month lead at the 1957 peak; by my nearly ex ante scoring the CLI misses the 1957 peak entirely. For the 1948 peak and the 1970 trough, I score leads two and three months longer than the official leads. For the other seven turns that can be compared, my leads all fall about three months short of the official leads. The mini-peak of 1966 is so clearly defined in the CLI series that there can be no doubt that the NBER—had it defined a December peak—would have recorded a six- to seven-month lead for the CLI. A general conclusion that emerges, therefore, is *a distinct tendency for the hindsight leads to exceed the foresight leads*. The record as established in Table 4 is not bad, but it is significantly inferior to the impression given by Table 3.

Neither Table 3 nor Table 4, however, provides any information about a possible tendency of the CLI to provide false signals.[25] Table 5 characterizes the entire set of monthly forecasts for the period 1948–71 in a two-way table that compares the predicted and realized results. First, it is immediately obvious that the predictions and realizations can by no stretch of the imagination be considered independent or unrelated to each other. Fully 82 percent of the observations lie on the main diagonal of the two-way table of predicted and realized results; only 50 out of 285 observations, less than 18 percent, represent errors of one kind or another. But clearly, independence is not a very interesting hypothesis to be testing. Of 285

25. Many earlier studies have been critical of the indicator approach on the grounds of an excessive tendency to signal false turns. The classic study that reaches this conclusion is Alexander, "Rate of Change Approaches."

Table 5. Turning Point Test of the Composite Leading Indicator, 1948–71[a]

Predicted result	*Realized result* No change	Peak	Trough	Total observations	*Percent of predicted turns that are false*[b]
No change (NC)	181	14	8	203	...
Peak (P)	28	29	0	57	49.1
Trough (T)	0	0	25	25	0
Total observations	209	43	33	285	...
Percent of realized turns that are missed[c]	...	32.6	24.2	...	...

Source: Calculations by author.

a. To test independence: obs. $\chi^2 = 276.2$; $\chi^2_{4;\,0.01} = 13.3$, where obs. χ^2 is the chi-square value for the observed data computed under the assumption that predictions and realizations are independent, and $\chi^2_{4;\,0.01}$ is the value that an observed chi-square with 4 degrees of freedom would exceed with probability 0.01 *if* predictions and realizations were independent.

b. A P (T) prediction is false if it occurs for a month other than the peak (trough) month, the three (two) months prior to the peak (trough) month, or the two months following the peak (trough) month, or if it occurs outside of a consecutive string of P (T) predictions from the peak (trough) month back to at most nine (six) months before the peak (trough) month.

c. A realized peak (trough) has been missed if the peak (trough) was not predicted for the peak (trough) month, the three (two) months prior to the peak (trough) month, or the two months following the peak (trough) month.

months, only 64 require a prediction other than no change (NC).[26] Thus a "naive" forecaster who never said anything but NC would have made 64 errors, an error rate of 22 percent, compared with 50 errors, an error rate of 18 percent, resulting from use of the CLI. The additional statistics shown in the last column and last line of Table 5 offer a useful way of viewing the 50 errors committed by the CLI. Nearly one-fourth (24.2 percent) of the months that should have carried a trough prediction received an erroneous NC prediction; nearly one-third of the months that should have carried a peak prediction received an erroneous NC prediction. But these are not the most serious errors; they point primarily to the fact that the turns are predicted with shorter leads (and somewhat less consistency) than one might like. The really serious error involves the false peak predictions. Of all the P predictions, nearly half (49.1 percent) were false. When the CLI signals a peak, flip a coin! And that is a serious indictment of the indicator.[27]

26. 6 Ps for each of the six peaks, 3 Ts for the 1967 trough, and 5 Ts for each of the other five troughs.

27. In *Signals of Recession* (pp. 108–14), Shiskin discusses the problem of false signals over the 1948–61 period. He attributes the false peak signals that he (and I) observe in 1951 to the pattern of the defense build-up due to the Korean War, finds it difficult to explain those of 1956, and rationalizes those in mid-1959 as due to the steel strike. I have already dealt with the latter by disregarding the mid-1959 signals, but I am not convinced that the 1951 case would have been discounted at the time of its occurrence.

But perhaps this conclusion is overly negative. By my test a "false" peak signal is one that is not followed directly by an economic recession of the classical sort. Some—Ilse Mintz especially—have argued that one ought to expand the notion of a recession to include a period of substantial decline in the rate of growth of the economy.[28] If this notion is accepted, a peak signal is not to be considered false if it is followed by a significant sustained slowdown in economic activity. While I continue to believe that a valid distinction can be drawn between a slowdown and a classical recession, I do not deny that my conclusion regarding false peaks would have to be softened considerably if the peak signals in question turned out to be reliable predictors of significant slowdowns in economic growth. Table 6 contains a chronology of the rate of economic growth immediately before and after the dates at which my test of the CLI records false peak signals. I have measured the growth rate by using the deflated CCI, but an alternative calculation based on the index of industrial production yields the same sort of results. Combining the evidence in Table 6 with the results of the earlier analysis of the CLI, I conclude that the April 1951 signal can easily be rescued from the false-signal designation. The November 1950 signal could have been regarded as useful if it had been maintained, but it was reversed for three months beginning in January 1951. Similarly, the February 1957 signal would have constituted a reasonable lead for the July 1957 peak, but it was reversed in April 1957. The signal in December 1959 was followed immediately by an up-tick in an already meager growth rate and the signal was reversed three months before the May 1960 peak, just as the level of the CCI was already heading downward. The other signals shown in Table 6 were soon followed by increases in the growth rate or by declines that were reversed before reaching significant proportions. Thus I conclude that seven of the eight signals shown in Table 6 must be regarded as false; the concept of a growth recession is not sufficient to rescue the CLI from the charge of excessive signaling of false peaks. Mintz's conclusion is quite the opposite of mine, and I submit that our differences result from the fact that her conclusion rests on an ex post interpretation of the CLI while mine continues to be based on an ex ante reading of the data. What Mintz's chronology can ignore as meaningless wriggles I read as confusing data that negate the value of the preceding signal.[29]

28. See Ilse Mintz, "Dating American Growth Cycles," in *Business Cycle Today*.

29. Lest anyone think that the CLI fails entirely to filter out false signals, let me cite the case of the index of 500 common stock prices. In some quarters, the stock market is

Table 6. Economic Growth Rates around Dates of False Peak Signals of the Composite Leading Indicator, 1948–71

	Annual rate of growth of the CCI (percent)[a]		
Date of false peak signal	*6 months prior to peak signal*	*First 3 months following signal*	*Second 3 months following signal*
November 1950	+21.7	+19.5	+8.0
April 1951	+12.5	−1.9	+1.0
May 1952	+7.7	+6.5	+27.6
March 1956	+6.3	+2.3	+7.6
June 1956	+2.3	+7.6	+6.0
February 1957	+7.9	−4.2	+2.2
December 1959	+1.4	+3.4	−1.3
July 1962	+8.7	+3.7	+4.3

Source: Calculations by author.
a. CCI = composite coincident indicator (the reverse trend adjusted CLI).

Up to this point, I have been charitable to the CLI by conducting ex ante turning point tests on historically revised data. In point of fact, what the forecaster has available to him in month t and what the newspapers print is a highly preliminary estimate of the CLI for month $t - 1$ and a first-revision estimate of the CLI for month $t - 2$. Indeed, the preliminary estimate is generally based on some six to eight components, rather than the full twelve. A true ex ante test of the CLI must be based on the preliminary and first-revision data, not the historically revised data. Obviously, it is not possible to obtain the preliminary and first-revision data for the 1966 short list CLI for a long enough period to run a useful test. I have, however, been able to gather such data for all the components back to 1961 and for many of them back to 1956. With these I was able to construct a pseudo-series of preliminary and first-revision CLI estimates.[30] My preliminary estimates are

regarded as a reliable cyclical indicator in and of itself. I have taken the series of monthly changes in the stock price index, made a reverse trend adjustment in it, and tested it in exactly the same way as the CLI was tested in Table 5. The results are striking. The stock market index yields a false-peak rate of 69.3 percent compared with 49.1 percent for the CLI! There are fifty-two false peak predictions and only twenty-three appropriate peak predictions for the stock market index, compared with twenty-eight false and twenty-nine appropriate peak predictions for the CLI.

30. I did *not* carry out the complex procedure corresponding to the sketchy outline provided above (pp. 343–48). Instead, I used the historically revised data to run a regression of the monthly changes in the CLI on the monthly changes in its components (excluding profits, for which monthly changes are not meaningful). This regression provided a very tight fit and is not subject to the criticism surrounding Table 2 because the CLI is

Table 7. Turning Point Test of the Composite Leading Indicator, 1956–71[a]

Predicted result	*Realized result*: *No change*	*Peak*	*Trough*	*Total observations*	*Percent of predicted turns that are false*[b]
No change (NC)	131	14	8	153	...
Peak (P)	13	13	0	26	50.0
Trough (T)	0	0	11	11	0
Total observations	144	27	19	190	...
Percent of realized turns that are missed[c]	...	51.9	42.1	...	...

Source: Calculations by author.

a. To test independence: obs. $\chi^2 = 135.4$; $\chi^2_{4;\ 0.01} = 13.3$, where the symbols are defined as in Table 5, note a.

b. See Table 5, note b.

c. See Table 5, note c.

always based on eleven components (see note 30) and should therefore be at least as reliable as the preliminary estimates of the kind now regularly published in *BCD*. In testing these data, I have applied exactly the same rules as previously, except that the two consecutive CLI changes to be considered for month t are the preliminary for t and the first revision for $t - 1$. In other words, an incorrect preliminary sign for $t - 1$ will not affect the prediction for month $t + 1$ if the first revision corrects the preliminary error.

I am able to conduct a turning point test of these data only for the period since 1956. To provide a valid comparison, Table 7 gives the results for the historical CLI for the 1956–71 period; the results for the preliminary-first revision CLI are contained in Table 8. Note first that the historical CLI itself fares less well in the 1956–71 subperiod than in the full period 1948–71.

in fact constructed as a linear function of its components. I then calculated two error series for each leading indicator component: one representing the difference between the preliminary and the historically revised change for each month; the other representing the difference between the first revision and the historically revised change for each month. I then used the coefficient estimates from the regression equation to build the preliminary errors into the historically revised ΔCLI series, yielding a preliminary ΔCLI series. Finally, I used the coefficient estimates to build the first-revision errors into the historically revised series. For 1956–60 I used preliminary and first-revision data for those series for which I had them, and the historically revised data for the rest. I would think that this biases the results in favor of the preliminary-first revision CLI, but it is vaguely possible that I have eliminated some offsetting errors and therefore biased the results against the preliminary data.

Table 8. Turning Point Test of Preliminary-First Revision Composite Leading Indicator, 1956–71[a]

	Realized result				
Predicted result	*No change*	*Peak*	*Trough*	*Total observations*	*Percent of predicted turns that are false[b]*
No change (NC)	130	9	8	147	...
Peak (P)	13	19	0	32	40.6
Trough (T)	0	0	11	11	0
Total observations	143	28	19	190	...
Percent of realized turns that are missed[c]	...	32.1	42.1	...	...

Source: Calculations by author.
a. To test independence: obs. $\chi^2 = 164.7$; $\chi^2_{4;\ 0.01} = 13.3$, where symbols are defined as in Table 5, note a.
b. See Table 5, note b.
c. See Table 5, note c.

The false-peak rate is still 50 percent but the missed turns are considerably more numerous (relatively). The most amazing result, however, comes from the comparison of Tables 7 and 8. The preliminary-first revision CLI outperforms the historically revised CLI. The margin of difference is not great —the false-peak rate is down to "only" 40.6 percent—but it is all one-sided. Table 9 provides further evidence in terms of the leads at turning points. Compared with the comparable turns in Table 4, no leads have shortened, and one, that at the 1969 peak, is longer by five months.

I had fully expected that the preliminary-first revision data would prove inferior to the revised data. In retrospect I see the obvious reason why my initial expectations were wrong. The entire procedure being evaluated depends only on the sign of a change in the CLI, not its magnitude. This is a weakness in the sense that no attempt is being made to extract quantitative

Table 9. Unofficial Leads and Lags of the Preliminary-First Revision Composite Leading Indicator at Cyclical Turning Points, 1956–71

Peak	*Lead (−) or lag (+) (months)*	*Trough*	*Lead (−) or lag (+) (months)*
July 1957	missed	April 1958	+1
May 1960	−1	February 1961	+1
December 1966	−6	March 1967	+2
November 1969	−4	November 1970	−3

Source: Calculations by author.

information from the CLI; but it becomes a strength when the forecaster is at the mercy of preliminary data. While the preliminary and first-revision CLI contain many quantitatively large errors, there are only 16 (out of a possible 190) incorrect preliminary signs, 8 of which are correctly revised a month later. In no case are two consecutive preliminary signs incorrect, and in only two cases are both a preliminary and the preceding first-revision signs incorrect.[31]

In closing this section, I hasten to reiterate that, though the preliminary-first revision CLI outperformed a poor competitor, it by no means performed so well as to confirm that a reliable cyclical indicator was in hand.

Spectral Analysis of the Leading Indicators

In this section I shall attempt to deal with two factors already noted. The first is the tendency of the CLI, as now constructed, to produce an excessive number of false peak signals. The second is the fact, noted in connection with Table 1, that the twelve series of leading indicators do not coincide with each other in their median lead times, and that each exhibits extreme variability around its own median lead time. A technique that can be employed most profitably to shed light on these matters is spectral analysis, a statistical procedure for ferreting out the cyclical components of a time series.[32]

Roughly speaking, spectral analysis views a time series as the sum of a number of other time series, each of which is a periodic series with a specific amplitude. Thus a given time series, A, may be thought of as the sum of the series A_1, A_2, and A_3, where A_1 is a series that exhibits large amplitude cycles with a five-year periodicity, A_2 is a series that exhibits small amplitude cycles with a two-year periodicity, and A_3 is a series that exhibits very small amplitude cycles with a two-quarter periodicity. One could then regard the series A as one with a major cycle component (A_1), a minor cycle component (A_2), and a very short cycle component (A_3). If A_1, A_2, and A_3 are

31. I also believe, as discussed in note 30 and subsequently, that my test was to some extent biased in favor of the preliminary-first revision test of the CLI.

32. Some useful references on spectral analysis are Granger, in association with Hatanaka, *Spectral Analysis;* George S. Fishman, *Spectral Methods in Econometrics* (Harvard University Press, 1969); and Marc Nerlove, "Spectral Analysis of Seasonal Adjustment Procedures," *Econometrica*, Vol. 32 (July 1964), pp. 241–86.

random time series (or stochastic processes), then the periodicity and amplitude of each is an approximate or average value and A_3 is the random noise component of the series A. Empirical spectral analysis succeeds in decomposing virtually any observed time series into a set of approximately periodic components that sum to the original time series,[33] and provides a measure, referred to as the "power," of how much of the variance of the original time series is attributable to each of the approximately periodic components. If a time series has been decomposed into a twenty-period cycle and a five-period cycle with powers of 5 and 10, respectively, then the short-period cycle accounts (empirically) for twice as much of the variance of the original time series as does the long-period cycle. The periodicities into which a time series is decomposed, along with the power measure at each periodicity, is known as the "spectrum" or "power spectrum" of the time series.

One may also compute a cross-spectrum between two time series. In this case each time series is decomposed into components of the same approximate periodicity—say, three components with approximate periodicities of twenty months, eight months, and five months. The cross-spectrum calculation provides two important measures for each periodicity: the "squared coherence," which is essentially the square of the correlation between a periodic component of one series and the like component of the other; and a "phase" displacement, which is the approximate number of time units (months, for example) separating the peaks (or troughs) of a periodic component of one series and the like periodic component of the other.[34]

The proper application of spectral analysis does not require that a time series actually be generated by purely cyclical phenomena, in the sense of an endogenous Metzlerian inventory cycle superimposed upon an endogenous Hicksian business fixed investment cycle.[35] I am not especially fond of any theory proposing an endogenous, self-perpetuating business cycle.

33. Just as least squares will provide the best fit between two variables whether they have anything to do with each other or not.

34. If the eight-month periodicity components of the two time series have a phase displacement of two months (one leads the other by two months), the cross-spectrum calculation "aligns" the series *before* computing the squared coherence. Thus a squared coherence of 0.90 would mean that, with proper aligning of the components, 90 percent of the variance of one approximate eight-month cycle can be explained by the other approximate eight-month cycle.

35. Of course, if the time series *were* to be so generated, the spectral decomposition *would* factor into those cycles.

Table 10. Relative Power Spectra for Indicator Variables

Forty-month periodicity = 1.00

	Change in indicator												
Periodicity (months)	*Deflated composite index of coincident indicators*	*Composite index of leading indicators*	*Workweek*	*Unemployment claims*	*Durable orders*	*Plant and equipment contracts*	*Business formation*	*Price-cost ratio*	*Stock prices*	*Materials prices*	*Building permits*	*Inventory change*	*Installment debt*
60	0.73	0.67	0.53	0.67	0.69	0.84	0.97	0.75	0.72	0.82	0.59	0.41	0.60
40	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
30	0.93	0.81	1.17	1.27	1.11	1.24	0.80	1.20	0.85	0.82	1.07	1.80	0.77
24	0.60	0.52	0.78	0.92	0.85	1.12	0.50	0.89	0.84	0.60	0.69	1.39	0.37
20	0.31	0.28	0.33	0.38	0.38	1.00	0.25	0.51	0.66	0.35	0.35	0.73	0.22
12	0.12	0.14	0.21	0.44	0.26	1.41	0.06	0.14	0.61	0.46	0.13	0.33	0.51
8	0.16	0.12	0.35	0.55	0.56	1.11	0.12	0.61	0.37	0.16	0.36	1.04	0.58
6	0.14	0.09	0.34	0.66	0.75	1.09	0.12	0.62	0.50	0.11	0.60	1.44	0.55
4.8	0.17	0.09	0.58	0.71	0.84	1.05	0.13	0.62	0.58	0.10	0.68	2.35	0.48
4	0.17	0.10	0.69	0.66	0.90	1.24	0.16	0.50	0.44	0.09	0.52	3.49	0.91
3.4	0.15	0.11	0.90	0.75	1.19	1.63	0.13	0.41	0.34	0.06	0.40	4.02	1.19
3	0.15	0.10	1.05	0.81	1.50	2.06	0.12	0.44	0.39	0.05	0.55	3.95	1.12
2.7	0.13	0.11	0.87	0.84	1.56	2.29	0.14	0.39	0.38	0.04	0.90	4.49	0.94
2.4	0.13	0.12	0.77	0.81	1.18	2.00	0.14	0.30	0.27	0.04	1.09	6.41	0.61
2.2	0.14	0.13	0.90	0.72	0.91	1.33	0.22	0.35	0.22	0.05	0.84	7.55	0.71
2	0.14	0.13	0.92	0.53	0.85	1.11	0.29	0.42	0.19	0.05	0.61	7.48	0.74

Source: Calculated by author from monthly first differences of the time series for the period from 1948 to mid-1972.

But casual empiricism and the behavior of econometric models do offer convincing evidence that the economy responds to various stimuli in a damped cyclic fashion; that, depending on the nature of the particular stimulus, this cyclical response mechanism may exhibit a number of waves of long or short period and of large or small amplitude before dying out; and that in any period of time the economy may well be responding to several stimuli *simultaneously*. The decomposition of an aggregative time series into components with perhaps two- to five-year periodicities is—in this view—a meaningful way to inquire into what is generally and inaccurately known as "the business cycle."[36]

Table 10 contains the power spectra calculated for the deflated CCI, the reverse trend adjusted CLI, and each of the components of the CLI (except for corporate profits). In each case the spectrum was calculated for the monthly first differences of the time series in question over the period 1948 through mid-1972. Since the power itself is in different units for different series, I have expressed the spectra in relative terms, showing for each periodicity of a given series its power relative to the power of the forty-month periodicity.[37]

In the power spectrum for the CCI, the twenty-four- to sixty-month periodicities clearly dominate. In other words, approximate cycles of two- to five-year periodicity account for far more of the variance of the CCI than do any cycles of shorter periodicity.[38] The CLI, constructed so as to signal the major cyclical movements in the CCI, displays a very similar spectral pattern involving clear dominance of the twenty-four- to sixty-month periodicities. The shorter periodicities in the range of two to eight months can hardly be interpreted as components of the business cycles that are the

36. One might well wish to consider a periodicity of about eight years—the classic Juglar cycle—but here a statistical problem arises. In a time series of about 300 months (1948–72), one could observe at most three eight-year cycles; that is simply too few for spectral techniques to produce a reliable estimate of a component with an eight-year periodicity.

37. The forty-month periodicity is a convenient base since most of the series in question exhibit their maximum long-periodicity power at forty months, which is surely relevant for the analysis of business cycles.

38. The first difference transformation affords a more reliable look at the spectrum but may have the effect of slightly exaggerating the importance of the forty- and thirty-month periodicities relative to the sixty-month periodicity. It is important to consider a rather fine breakdown of periodicities in the two- to twelve-month range because it is the flatness of the spectrum over such a range that permits the conclusion that the short-periodicity "cycles" can be regarded as random noise events.

focus of this paper. Rather, they correspond to what must be regarded as short-period disturbances, minor spurts and stalls resulting from any number of events that continually intrude on the underlying course of the economy. They are not, however, trivial with respect to the performance of the CLI. They are present in both the CCI and the CLI and correspond to the occasional reversals—or short string of reversals—that occur in each of the series during an otherwise unidirectional movement; that is to say, they correspond to the false signals often displayed by the CLI. The CLI has, in a sense, reproduced the CCI too faithfully. While showing the same behavior as the CCI with respect to the periodicities of interest, it is also duplicating the tendency to exhibit aberrant phase reversals that give the same signal as a true business cycle phenomenon; the CLI does not succeed in filtering out all movements that are pure noise from the point of view of the reference cycle.

An inspection of the spectra for the short list variables yields some clues to the noise phenomenon. Every one of the leading indicator variables exhibits a relative peak in its power spectrum at either the thirty- or forty-month periodicity, which, of course, is precisely what makes each a potential cyclical indicator of some sort.[39] But contrast business formation and inventory change. By itself, business formation has a spectrum very much like that of the CCI: substantial power in the long periodicities, and much less power in the noise range. Inventory change, on the other hand, is totally dominated by noise—by brief up and down spurts that account for far more variance than do the longer cycles. Materials prices has a nearly ideal spectrum for a leading indicator: dominant power in the long periodicities, and little noise. Unfortunately, its long-period cycles are not well correlated with those of the CCI and its lead times leave much to be desired. Most of the other leading series fall somewhere in between business formation and inventory change in their spectral patterns, but with a tendency toward much noise. In view of the spectral contrast between the CLI and its components, the noise elements in the component series apparently are not highly correlated, while the long-period elements are. The averaging process therefore results in the CLI exhibiting relatively little noise, but still too much.

Table 11 contains the cross-spectrum results for the CCI and CLI series.

39. I call the data in Table 10 "potential" indicators because they give no information about leads, lags, conformity, and so on.

Table 11. Cross-Spectral Statistics for the Change in the Composite Coincident Indicator vs. the Composite Leading Indicator

Periodicity (months)	*Squared coherence*	*Phase lead (−) or lag (+) of CLI relative to CCI (months)*
60	0.85	−7.6
40	0.92	−4.3
30	0.93	−2.8
24	0.83	−2.3
20	0.54	−1.5
12	0.46	+0.1
8	0.49	−0.2
6	0.35	−0.2
4.8	0.19	−0.4
4	0.14	−0.3
3.4	0.18	0.0
3	0.19	+0.2
2.7	0.11	0.0
2.4	0.10	0.0
2.2	0.16	+0.1
2	0.10	+0.1

Source: Calculated by author from time series for the period from 1948 to mid-1972.

Three conclusions stand out. First, the long-period cycles in the two series are very highly correlated: the CLI is potentially a good predictor of the major swings exhibited by the CCI. Second, the CLI clearly leads the CCI in the major cycles, but by a highly variable period, extending from more than seven and one-half months in the five-year cycle component to less than two and one-half months in the two-year cycle component. The implication is that, *even if all goes well,* the lead time of the CLI will be highly unreliable, depending on what kind of cyclical response mechanism happens to be dominating the business cycle turn. The third point involves the poor correlation in the noise range: a false signal of the CLI need not even correspond to an impending aberration in the path of the CCI. I submit that it would be highly desirable if the CLI exhibited less noise and if its lead time were substantially more uniform in the major cycle components.[40]

40. It might be argued that long notice is more important in the case of an impending major downturn than a minor downturn. But that is not the message of Table 11. Remember that all the postwar downturns have been short, and 1957–58, the sharpest postwar recession, was among the shortest. Also remember that all of the periodicities operate simultaneously, and to justify a desire for the lead time to vary with the magnitude of the impending phase requires pinpointing of the major cause of the turn ahead of time.

Table 12. Cross-Spectral Statistics for the Change in the Composite Coincident Indicator vs. the Leading Indicators

Periodicity (months)	Change in indicator[a]											
	C^2	(−) or (+)	C^2	(−) or (+)	C^2	(−) or (+)	C^2	(−) or (+)	C^2	(−) or (+)	C^2	(−) or (+)
	Workweek		*Unemployment claims*		*Durable orders*		*Plant and equipment contracts*		*Business formation*		*Price-cost ratio*	
60	0.78	−6.5	0.80	−3.9	0.91	−6.6	0.07	−34.5 or +25.5[b]	0.48	−4.4	0.40	−9.7
40	0.88	−3.8	0.86	−3.4	0.91	−3.7	0.11	−10.1	0.78	−5.0	0.33	−3.1
30	0.94	−2.7	0.91	−2.5	0.85	−2.3	0.19	−4.2	0.79	−4.5	0.63	−1.2
24	0.98	−2.0	0.88	−1.9	0.79	−2.2	0.11	−2.4	0.67	−4.2	0.75	−1.3
20	0.89	−1.6	0.63	−1.2	0.53	−2.3	0.00	−7.8	0.40	−2.6	0.45	−0.5
6	0.20	−0.3	0.24	−0.0	0.14	−0.6	0.02	−1.8	0.01	+0.5	0.30	−0.2
4	0.01	+0.2	0.05	+0.3	0.06	−0.3	0.02	−1.1	0.02	+1.6	0.10	−0.1
3	0.02	+0.2	0.09	+0.0	0.04	−0.0	0.02	+0.6	0.09	+0.3	0.05	−0.2
2	0.00	−0.6	0.18	+0.0	0.16	+0.1	0.05	+0.9	0.01	−0.3	0.09	+0.2
	Stock prices		*Materials prices*		*Building permits*		*Inventory change*		*Installment debt*			
60	0.66	−9.8	0.25	−6.0	0.63	−20.5	0.59	−6.0	0.56	−8.3		
40	0.67	−8.1	0.20	+0.5	0.71	−11.4	0.63	−1.8	0.71	−5.1		
30	0.47	−6.3	0.46	+1.7	0.79	−8.0	0.72	−1.1	0.63	−3.9		
24	0.15	−5.0	0.37	−0.1	0.66	−6.4	0.64	−1.5	0.61	−1.9		
20	0.08	−3.9	0.01	−2.9	0.16	−4.9	0.37	−2.2	0.61	+1.0		
6	0.01	+0.3	0.00	+0.0	0.05	−1.0	0.14	−0.4	0.08	+1.0		
4	0.03	−0.3	0.03	+0.1	0.04	−0.2	0.06	−0.2	0.02	+1.3		
3	0.16	+0.8	0.01	+0.4	0.02	+0.4	0.02	−0.0	0.09	−0.1		
2	0.01	+0.0	0.00	−0.2	0.03	−0.1	0.08	+0.4	0.03	+0.1		

Source: Calculated by author from the time series for the period from 1948 to mid-1972.
a. C^2 = squared coherence; (−) and (+) = phase lead and lag, respectively, of indicators relative to CCI, in months.
b. See text note 41 and accompanying discussion.

Underlying the variable lead time exhibited by the CLI is the fact that it is constructed by a contemporaneous averaging of a dozen series, each of which exhibits a variable lead time pattern. I have already pointed this out in connection with Table 1. It is even more graphically displayed in Table 12, which contains the cross-spectrum statistics for the CCI compared with each of the leading indicator series whose power spectrum is shown in Table 10. In Table 12, I have shown only the major cycle periodicities and a selection of the noise components; the format is the same as in Table 11, showing first the squared coherence (C^2) and then the phase lead or lag, one series at a time. The average workweek in manufacturing, initial claims for unemployment insurance, and new orders for durable goods stand out with the highest degree of explanatory power in the long periodicities, but their lead times are quite short for the thirty- and twenty-four-month components. In the case of the workweek, for example, the lead time ranges from six and one-half months in the five-year cycle down to only two months in the two-year cycle. Materials prices, which had a nearly ideal spectral pattern, can be classed as a leading series only for the very longest cycle and has relatively low explanatory power, except possibly for the thirty-month cycle, where it *lags* the CCI by nearly two months. Plant and equipment contracts exhibits quite long lead times but has surprisingly low coherence.[41] Net business formation stands alone as the indicator that has both reasonably high coherence *and* a uniform lead time throughout the long periodicities.[42] The unmistakable modal result is reasonably high coherence with substantial variability in lead time.

The data in Tables 10 to 12 provide important information about the properties of the indicator series and the results obtained in the turning point tests conducted earlier. This information can be used in an attempt to improve the results.

SLI: An Alternative to the CLI

I shall use the results discussed in previous sections to suggest an alternative technique for the construction of a cyclical indicator. I maintain the basic assumption, discussed earlier, that in the construction of an indicator,

41. Its lead time in the five-year cycle is so close to half the cycle length that it is not clear whether it should be considered a long lead or a long lag; either could be rationalized.

42. It is interesting to note, in this connection, that net business formation is the very weakest variable in the multiple regression equation reported in Table 2.

it is desirable to combine many potentially duplicative signals of the same impending event. But, in contrast to the NBER time-domain technique, I shall focus on the cyclical components that are to be predicted. Specifically, I shall use the spectral and cross-spectral statistics to construct one indicator component corresponding to each of the four major cycle periodicities that have been discussed: sixty, forty, thirty, and twenty-four months. In the construction of each of these components, I shall attempt to highlight the corresponding cycle while minimizing the noise content. I shall then use the information in the CCI spectrum to combine the four indicator components into a single indicator, which I shall refer to as the "spectral leading indicator" (SLI).

The construction of the SLI is best illustrated by an example using the hypothetical data in Table 13. Two series, X and Z, are considered to be leading indicators for the series Y. Periodicity p_4 is a noise component, while p_1, p_2, and p_3 correspond to the "real cycles" the SLI is to predict. Assume, for ease of exposition, that all three series have comparable units of measurement and have trivial power at p_4.

The indicator component for periodicity p_1, say I_1, is constructed as

$$(2) \qquad I_1(t) = \frac{\sqrt{0.8}\ X(t-5) + \sqrt{0.5}\ Z(t-7)}{\sqrt{0.8} + \sqrt{0.5}}.$$

Since X leads Y by five months at p_1, and Z leads Y by seven months at p_1, I average X and Z with the indicated lags. The weights in the average are the coherences ($\sqrt{C^2}$), so that X carries appropriately more weight than Z

Table 13. Hypothetical Spectral and Cross-Spectral Statistics, Three Series

	Series Y	Series X		Series Z	
Periodicity	Power	Squared coherence	Phase lead (−) or lag (+) (months)	Squared coherence	Phase lead (−) or lag (+) (months)
p_1	2	0.8	−5	0.5	−7
p_2	3	0.6	−2	0.9	−5
p_3	2	0.9	+1	0.6	−5
p_4	0.01	0.1	−1	0.2	−1

Source: Hypothetical data constructed by author. See discussion in text.

in predicting the p_1 component of Y. The indicator component for periodicity p_2 is similarly defined:

$$I_2(t) = \frac{\sqrt{0.6}\ X(t-2) + \sqrt{0.9}\ Z(t-5)}{\sqrt{0.6} + \sqrt{0.9}}. \tag{3}$$

In this case, Z receives the higher weight. In the case of the p_3 indicator the X variable must be disregarded, despite its high explanatory power, since it is a lagging indicator at p_3. Thus:

$$I_3(t) = Z(t-5). \tag{4}$$

Note that $Z(t-5)$ has been used twice, in I_2 and again in I_3.

In combining I_1, I_2, and I_3 into a single indicator, I use the power of the Y series, which indicates that the p_2 periodicity is of greater importance than either p_1 or p_3. Thus:

$$SLI(t) = \frac{\sqrt{2}\ I_1(t) + \sqrt{3}\ I_2(t) + \sqrt{2}\ I_3(t)}{\sqrt{2} + \sqrt{3} + \sqrt{2}}. \tag{5}$$

In averaging the components to obtain the SLI, the square roots of the power measures are used since the power represents a variance decomposition. Note that the SLI, as constructed in this example, becomes "available" two months prior to the date to which it refers.

The hypothetical example fails to bring out three points that are of great relevance for the practical application of the method to the leading indicators. The first is that the short list indicators do not have comparable units of measurement. I have handled this in essentially the same fashion that the NBER employs in the construction of the CLI: the monthly changes in each leading indicator series have been normalized by dividing through by the average absolute monthly change over the sample. The second is that the power of the noise components of the leading indicator series is not at all trivial. I have treated this problem as follows: if the cross-spectrum indicates that the change in BCD_i leads that in CCI by, say, seven months at a particular frequency, and if the spectrum of change in BCD_i exhibits substantial noise (which is true of all of the series except for the changes in business formation, stock prices, and materials prices), then I use

$$\frac{\Delta BCD_i(t-6) + \Delta BCD_i(t-7) + \Delta BCD_i(t-8)}{3}$$

instead of $\Delta BCD_i(t-7)$ itself in the construction of the indicator com-

Table 14. Spectral and Cross-Spectral Statistics for the Spectral Leading Indicator and the Change in the Composite Coincident Indicator

		Cross spectrum	
Periodicity (months)	*SLI spectrum*	*Squared coherence*	*Phase lead (−) or lag (+) of SLI relative to CCI (months)*
60	0.62	0.81	−3.2
40	1.00	0.90	−2.1
30	1.01	0.91	−0.9
24	0.67	0.83	−0.6
20	0.26	0.50	+0.2
12	0.04	0.11	+1.1
8	0.05	0.35	+0.6
6	0.03	0.13	+0.4
4.8	0.02	0.06	+0.0
4	0.01	0.00	−0.3
3.4	0.01	0.04	+0.8
3	0.005	0.14	+0.4
2.7	0.005	0.10	−0.1
2.4	0.004	0.04	−0.6
2.2	0.003	0.00	−0.6
2	0.002	0.01	−0.7

Source: Constructed by author from spectral and cross-spectral statistics in Tables 11 and 12.

ponent.[43] Finally, I found—for the same reasons mentioned in connection with the CLI—that a reverse trend adjustment was needed. The final SLI, then, has exactly the same time trend as the change in the CCI.[44]

Table 14 contains the spectrum of the SLI and the cross-spectrum statistics for the SLI and the change in the CCI. The spectrum of the SLI displays very clear dominance of the long periodicities and an extremely low noise level. Further, the long-period cycles in the SLI have very high coherence with the corresponding components of the CCI, and the phasing is reasonably uniform in the twenty-four- to sixty-month components. The-

43. There are two exceptions to this averaging rule. For the twenty-four-month component I averaged only two monthly changes, for the workweek and unemployment claims. The indicated lead was two months in each case, and I did not wish to use any data not "available" at least two months in advance. I also disregarded leads of twenty months or more in the interest of not losing too many data points.

44. The SLI is itself a change indicator, not a levels indicator, since it is constructed from the monthly changes in the leading series.

oretically, the SLI should lead the CCI by two months since the SLI dated in month t is based on data available two months prior to month t. The phase results are not perfectly uniform in the twenty-four- to sixty-month components, but the average lead is 1.7 months and the range is only 2.6 months compared with a spread of 5.3 months in the cross-spectrum results for changes in the CLI and CCI (see Table 11).[45]

The results in Table 14 are reasonably encouraging. False peaks will probably be substantially less troublesome with the SLI than with the CLI. The major cyclical movements of the CCI ought to be very well defined by the SLI. However, the lead time is apt to be shorter than one might wish.

Turning Point Tests of the SLI

The turning point tests on the SLI are conducted under the same rules employed for the CLI, with one exception. In the case of the CLI, the prediction of a turning point required two consecutive negatives during an upswing or two consecutive positives during a downswing. The purpose of this requirement was to filter out false signals. The SLI, however, has been constructed so as to highlight the long-period cycles while minimizing the short-period noise. It is therefore reasonable to expect that additional filtering is unnecessary. In the case of the SLI, I count any negative during an upswing as a peak signal, and any positive during a downswing as a trough signal.[46]

Table 15 presents the leads and lags of the SLI for the turning points since 1949. The 1948 peak is not an available turning point for the SLI since thirteen observations are lost at the beginning of the sample period in the construction process. Those components of the SLI that become available the earliest do show declines right from the start in mid-1948, and the SLI itself quite likely would have led the 1948 peak by at least several months;

45. I am unhappy about the statistics for the twenty-month periodicity, and would like to have included a twenty-month component in the SLI. As Table 12 reveals, though, there simply are no long lead–high coherence results for the twenty-month periodicity. The cross-spectrum results for the SLI at the twenty-month cycle (and possibly the twenty-four-month cycle as well, which shows a lead of only 0.6 month) provide a clue that the SLI may not pick up troughs well whenever the downswing is dominated by the shorter of the long periodicities.

46. If a peak signal has occurred, it takes only one positive to define a false peak; a false trough signal is symmetrically defined.

Table 15. Leads and Lags of the Spectral Leading Indicator at Cyclical Turning Points, 1949–71

Peak	*Lead (−) or lag (+) (months)*	*Trough*	*Lead (−) or lag (+) (months)*
		October 1949	0
July 1953	−1	August 1954	−3
July 1957	−2	April 1958	+2
May 1960	0	February 1961	+2
December 1966	0	March 1967	+2
November 1969	0	November 1970	−3

Source: Same as Table 14.

but I have made no attempt to approximate the lead and have simply omitted the 1948 peak from Table 15. A comparison of Table 15 with Table 4 reveals a tendency for the SLI to exhibit somewhat shorter leads than the CLI. The differences, however, are exaggerated by the absence from Table 15 of the 1948 peak, at which the CLI exhibits its longest lead. The two major differences at peaks occur in 1957 and at the mini-peak in 1966. In the latter case, the CLI exhibits a six-month lead, while the SLI coincides at the peak. The SLI very clearly dominates the CLI at the 1957 peak. While the CLI misses the peak entirely, the SLI leads by two months. At troughs the SLI is never better than the CLI and performs as well only in the case of the 1967 and 1970 turns.

Comparison of the performance of the two indicators suggests that, while the differences are not great, the CLI has the clear edge in terms of lead time.[47] This loss of lead time cannot be worth bearing unless the SLI is substantially superior to the CLI in terms of its false-peak properties. Recall that the CLI provided twenty-eight false peak signals in the 1948–71 period (see Table 5). None of these occurred prior to 1949, and a comparison with the number of false peak signals for the SLI in the 1949–71 period is therefore relevant. Compared with twenty-eight false peak signals for the CLI, the SLI series records only seven, one-fourth as many.[48] The missing SLI data in 1948 make it impossible to present a two-way table equivalent to Table 5, but Table 16 presents the data on the SLI for the 1956–71 period

47. One would have to give huge weight to the 1957 peak to reverse this conclusion.

48. In fact, five of the seven false peak months arise as a single string of predictions starting in mid-1951 and containing a four-month period during which the deflated CCI remained essentially flat. This corresponds to the April 1951 peak prediction of the CLI which was discussed in connection with Table 6.

and is directly comparable to Table 7. The conclusions are immediately clear. The false-peak rate for the SLI is only 13.3 percent, compared with 50 percent for the CLI over the same period. The SLI exhibits the same number of correct peak predictions (thirteen) as the CLI, and misses fewer months when it should be giving a peak signal. The combined evidence from Tables 15 and 16 implies that the SLI is more consistent across peaks than the CLI, and is far superior with respect to false peak signals. Table 16 does reveal, however, that the SLI performs less well than the CLI at troughs. The CLI exhibits no false trough signals during 1956–71, the SLI shows two; the CLI failed to pick up eight of the nineteen months that should have carried a trough prediction, the SLI failed in eleven. This finding is merely a reaffirmation of the trough comparison made in connection with Table 15.

Compared to the CLI, the SLI exhibits an extremely low false-signal rate and is more consistent at the peaks, but at the cost of somewhat poorer lead-time performance. A substantially longer lead time might be worth a few more false signals. To this end, I have experimented with an alternate SLI that should provide longer leads at the turning points. The SLI was constructed by employing cross-spectral leads of two months or more. In the construction of the alternate SLI, I required cross-spectral leads of five months or more and thus disregarded a number of short leads that exhibit quite high coherence with the CCI. The leading indicator unemployment claims, for example, exhibits short leads and high coherence at the three longest business cycle periodicities: a 3.9-month lead at the sixty-month

Table 16. Turning Point Test of the Spectral Leading Indicator, 1956–71[a]

	Realized result				
Predicted result	*No change*	*Peak*	*Trough*	*Total observations*	*Percent of predicted turns that are false*[b]
No change (NC)	143	11	11	165	...
Peak (P)	2	13	0	15	13.3
Trough (T)	2	0	8	10	20.0
Total observations	147	24	19	190	...
Percent of realized turns that are missed[c]	...	45.8	57.9	...	...

Source: Calculations by author.
a. To test independence: obs. $\chi^2 = 137.1$; $\chi^2_{4;\,0.01} = 13.3$, where symbols are as defined in Table 5, note a.
b. See Table 5, note b.
c. See Table 5, note c.

Table 17. Leads and Lags of the Alternate Spectral Leading Indicator at Cyclical Turning Points, 1949–71

Peak	*Lead (−) or lag (+) (months)*	*Trough*	*Lead (−) or lag (+) (months)*
		October 1949	−1
July 1953	−1	August 1954	−6
July 1957	+2	April 1958	0
May 1960	−2	February 1961	−1
December 1966	−7	March 1967	0
November 1969	−3	November 1970	−1

Source: Calculations by author.

cycle ($C^2 = 0.80$), a 3.4-month lead at the forty-month cycle ($C^2 = 0.86$), and a 2.5-month lead at the twenty-four-month cycle (C^2 ⅝ 0.91). All of these leading signals are included in the corresponding components of the SLI; none is included in the alternate SLI. The alternate SLI is thus "available" as a cyclical indicator with three months more lead time than the SLI itself.[49] But it is inherently less reliable than the SLI and may therefore be expected to exhibit more false turn signals in addition to longer lead times.

The improvement in the lead times resulting from the use of the alternate SLI is striking and can be seen by comparing Table 17 with Table 15. The results shown in Table 17 are clearly superior to those of the SLI. At only two turning points—1957 and 1970—is the SLI superior; at all other turns the alternate SLI exhibits the same or longer leads, ranging up to seven months longer at peaks, and from one to three months longer at troughs. Considering all peaks and all troughs, the average lead is about two months longer for the alternate SLI over all turning points. In fact, the alternate SLI can be judged to be superior even to the CLI at the turns shown for both: at four out of five peaks, the leads are longer for the alternate SLI than for the CLI; and at four out of six troughs, the leads are at least as long for the alternate SLI as for the CLI.

The false-peak record is not as favorable for the alternate SLI as for the SLI itself, but it is superior to that for the CLI. During the 1949–71 period,

49. For example, the SLI constructed using data through January 1960 refers to March 1960; the alternate SLI constructed using data through January 1960 refers to June 1960.

Table 18. Turning Point Test of Alternate Spectral Leading Indicator, 1956–71[a]

Predicted result	Realized result: No change	Peak	Trough	Total observations	Percent of predicted turns that are false[b]
No change (NC)	132	9	4	145	...
Peak (P)	12	19	0	31	38.7
Trough (T)	0	0	14	14	0
Total observations	144	28	18	190	...
Percent of realized turns that are missed[c]	...	32.1	22.2	...	...

Source: Calculations by author.
a. To test independence: obs. $\chi^2 = 208.9$; $\chi^2_{4;\,0.01} = 13.3$, where symbols are as defined in Table 5, note a.
b. See Table 5, note b.
c. See Table 5, note c.

the alternate SLI exhibits twenty false peak predictions, nearly three times as many as the SLI itself, but eight fewer than the CLI.[50] Table 18 contains the two-way table of prediction results for the 1956–71 period. A comparison with Table 16 reveals that the alternate SLI is superior to the SLI in every statistic with the very clear exception of the false-peak rate. A comparison with Table 7 shows the alternate SLI to be slightly better than the CLI with respect to false peaks.[51]

It is indeed possible to trade off false signals and lead time. I have reported on one experiment that yields such a result. For reasons that I will try to make clear in the concluding section, I have not pursued the matter any further. Suffice it to say here that one can readily improve upon the CLI as an indicator of major turning points. This section has shown that a spectral leading indicator—based on no more raw data than the CLI employs—can be constructed to exhibit (a) at least as good leading behavior as the CLI at the recognized turns in economic activity, and (b) an appreciably better record with respect to false turn signals.

50. The prediction rules employed for the alternate SLI are identical to those for the SLI; no additional filtering was used despite the fact that the alternate SLI should be expected to be a good deal noisier than the SLI itself.

51. The 1956–71 period exaggerates the closeness of the false-peak rate for the alternate SLI and the CLI; over the longer, 1949–71, period the CLI was distinctly worse than the alternate SLI, as noted in the text.

Summary and Recommendations

In this study I have attempted to provide a fairly complete analysis of the performance of the existing composite index of leading indicators as a predictor of the major swings in economic activity in the period since the Second World War. The performance of the CLI leaves much room for improvement. Indeed, the conclusions of a fair number of earlier studies—some cited above—regarding excessive false signaling is heavily underscored here. I have found, on the basis of a set of rules that permit the CLI to perform reasonably well at recognized turning points, that it exhibits a 50 percent false-peak rate. Half of all the peak predictions given out by the CLI turn out to be false signals. I have also found that a reasonable calculation of ex ante lead times produces a substantially shorter set of leads than appears in the official data on the basis of ex post judgments.

On the more positive side, an experiment designed to reveal the differences in performance of the historically revised CLI and the preliminary-first revision CLI established virtually none of any significance. The test, by its nature, was to some extent biased against showing large differences, but there was still considerable latitude for differences to emerge and they simply did not. If the CLI is used as a qualitative predictor—that is, as a direction-of-change predictor—then predictions that employ the data as they become available in published sources will differ little from the ex post predictions resulting from subsequently revised data. This is a powerful finding, but its practical significance is diminished by the poor quality of the performance of the historically revised CLI itself.

The technique of spectral analysis was employed to shed light on the performance characteristics of the existing CLI. Two major findings emerged: a verification of the fact that the averaging process in the CLI construction is inadequate in filtering out noise (hence the high false-signal rate), and a clear indication that simple contemporaneous averaging of the twelve component leading indicators fails to make the most efficient use of the available signal information.

I have suggested an alternative calculation using the lead-lag behavior estimated by cross-spectral techniques and resulting in an indicator that I termed the spectral leading indicator. In the construction of the SLI, I employed the clues from the spectral results to highlight the major cyclical swings while minimizing the likelihood of false signals.

As first constructed, the SLI proved vastly superior to the CLI with respect to the presence of false signals, but was somewhat inferior to the CLI with respect to lead times at the recognized turning points in economic activity. An alternate SLI was constructed to exhibit a longer lead time, although at the likely cost of lesser reliability. The alternate SLI did in fact exhibit longer lead times than either the SLI itself or the CLI. Its false-peak rate proved to be somewhat better than that of the CLI (39 percent compared with 50 percent for the 1956–71 period, and 49 percent for the more inclusive 1948–71 period), but was distinctly inferior to that of the SLI itself.

The spectral construction of a leading indicator seems to me to be a procedure worthy of further investigation. While my results are highly suggestive, they do not prove the procedure beyond all doubt. But—and this is critically important—I have not given the technique its fairest possible chance. After all, my experiments were based strictly on the leading indicator components that other researchers have judged the most useful for the construction of the CLI. The spectral indicator, I submit, does outperform the CLI even on its own data base; but it can be expected to do considerably better if it is based on data more appropriate to its own purposes. It must be possible, by considering other potential leading series, to find some reliable leading indicators for the twenty-month cycle component; none exists among the 1966 short list. It is also likely to be possible to find some indicators to substitute for the large number of two- to four-month leads contained in the SLI as I was forced to construct it. Such a discovery might well permit construction of a spectral indicator that combines the false-peak rate of the SLI and the lead-time properties of the alternate SLI.

It may also be that the specific construction technique that I have suggested fails to make the most efficient use of the spectral information. This too calls for experimentation—and perhaps theory as well.

I conclude with two personal judgments. To seek to construct a reliable leading indicator would be eminently worthwhile. And the field is still wide open.

Comments and Discussion

Alan Greenspan: Saul Hymans has done an impressive job in extracting about as much information as I think one can get from what I consider a very weak data base. In fact, his paper almost succeeded in shaking my belief in the ultimate unusability of a composite leading indicator. But I do continue to have some very serious questions about the basic underlying procedure of such a composite indicator approach.

I don't think anyone questions that every individual item that serves as a leading indicator has forecasting value. At issue is whether some combination of these indicators, through a certain synergism, provides more information than the sum of what is available merely by examining the individual elements.

What has always disturbed me about this type of procedure is its ambiguity with respect to theory. On the one hand the process appears wholly empirical. Analysts throw into a hopper a huge number of measures of the economy and filter out those that happen to have a pattern of leading the reference cycles of the composite coincident indicator over the particular period chosen.

Yet clearly all is not empirical, since a number of things are left out. One —which is, incidentally, an excellent indicator of the stock market—is the length of women's skirts. If one were simply sifting through numbers, he would tend to use that. In fact, the particular short list of indicators in the CLI exhibits a heavy emphasis on investment incentives and actions.

Five of the twelve indicators reflect profits, stock prices, or capital goods. Hence, one can argue that at least some crude cyclical theory is implicit in the indicator approach. But unless one has a formal, theoretical view as to why the particular indicators that led in the past should lead in the future, the use of any one set to forecast future turning points is obviously on shaky conceptual ground.

What underlying theory can be inferred from the choice of indicators appears to be hidden in some sort of black box. The very heterogeneity of the short list of twelve indicators makes them very difficult to deal with. They are a mixture of trendless ratios, such as the workweek and price per unit of labor cost, and geometric trend series, such as corporate profits. The latter are a special problem when prices are rising rapidly, as they have been in the most recent period. Either the figures should be deflated, as Hymans points out in a footnote, or at least logarithmic first differences should be used. I am not sure that the trend adjustment applied to the total composite leading indicator or the spectral leading indicator appropriately compensates for this problem, especially for the current period, when prices have been rising so rapidly, clearly distorting some series relative to others.

While multicollinearity is obviously desirable in an indicator approach, I do not think it is appropriate to utilize series that contain overlapping information. For example, the series on new orders of durable goods industries and contracts for plant and equipment are both heavily weighted with the same new orders of manufacturers of capital goods equipment. Profits after tax adjusted for trend and price per unit of labor cost are conceptually close, and would show a much higher correlation were the data more accurate.

In evaluating the performance of these series, I question the use of discretionary ad hoc adjustments such as those Hymans makes for strikes. At what point does one stop? If special allowance is made for an auto strike or steel strike, why not Phase 4 or Phase 1½ or devaluation?

Still on the question of evaluating performance, Hymans uses the criterion that a leading indicator must give an indication *before* the event. Yet since even close observers rarely know as long as two months after the fact that the economy has turned down, anything that would confirm a peak just as it occurred, or shortly thereafter, also would have some value.

The most interesting thing in Hymans' analysis was the reliability of the series on net business formation as an indicator. Finding an indicator like this by spectral analysis is in itself a valuable result.

While I must admit I am basically skeptical about an aggregated indicator approach, I suspect that if it is to be salvaged as a useful tool and one that does not suffer from the problem of false signals, the approach that Hymans has taken is the right route.

Julius Shiskin: Saul Hymans' paper is a welcome contribution to this field. As he said, it has been a long time since the last review of the usefulness of this indicator.

My impression is that short-term forecasters do not use the leading indicator index, or the indicators themselves, in a mechanical way. For example, many people are asking Alan Greenspan's question: What is the relevance of the leading indicator index today in view of the fact that several of the component series are expressed in terms of highly inflated dollars? I think that's a healthy attitude, and I am very glad to see it.

I want to make three principal comments on this paper. My first comment is that I think Hymans' standards are too high. He is asking more from the composite leading indicator index than any of its advocates are asking. For example, in 1950, Geoffrey Moore wrote, "There is some ground for confidence that objective use of these methods will at least reduce the usual lag in recognizing revivals or recessions *that have already begun*." In 1955, he wrote, "It is important to be clear about what these results do *not* mean, as well as what they do mean. They do not mean that one can get much advance notice that a general business contraction is beginning or is coming to an end. They do help one to recognize these events at about the time they occur. Even then there is some risk of error."[1]

In April 1972, I wrote in an article for *The Washington Post*, "The difficulties of forecasting being what they are, the wise forecaster will take advantage of all the help he can get. He will take into account the complex effects of monetary and fiscal policy on the cyclical process, other important economic policy actions such as wage and price controls, and unusual developments such as big strikes and military events. He will make judgmental analyses of GNP accounts and use econometric models based on both Keynesian and monetary theories. And he will take advantage of contributions of the leading index and historical knowledge about business cycle behavior.

1. Geoffrey H. Moore, *Statistical Indicators of Cyclical Revivals and Recessions*, Occasional Paper 31 (National Bureau of Economic Research, 1950), p .76; Moore, "Leading and Confirming Indicators of General Business Changes," paper delivered at the Annual Midwest Conference on Business Indicators, 1955, and published in Moore (ed.), *Business Cycle Indicators* (Princeton University Press for the National Bureau of Economic Research, 1961), Vol. 1, p. 79.

"... What then can emerge is an informed judgment on the part of the forecaster on [where the economy has been, where it most] likely is, and what the rough probabilities are of its moving in various directions and at approximately what speeds."[2]

The essential question is whether the analyst can make better short-term forecasts by including the leading indicators and their index in his tool kit of data and forecasting techniques. It is clear from my quotations that I believe he can.

Second, I would view the question of false signals of cyclical turning points differently from the way Hymans does. This issue was popularized by Paul Samuelson, who once said, "Stock prices have accurately predicted nine of the last five recessions." In considering such criticisms, I believe the definitions of the National Bureau of Economic Research must be modernized to include slowdowns or retardations in economic growth, or growth recessions (I find either of these terms acceptable). Further, rapid inflation must be taken into account in addition to recessions.

I think it is most useful to consider growth cycles and NBER business cycles as a family of cyclical episodes ranging from the deep depression in 1933 to the mild slowdown in 1967. All of these must be taken into account in setting a chronology of cycles, and in determining whether leading indicators in fact lead.

In this context, what differences in economic policy are appropriate for a contraction that just misses the recession definition compared with one that just meets it? For illustrative purposes, assume that an index of aggregate economic activity of 70 is the demarcation line between "recession" and "no recession." A planned slowdown from a level of 110 (to cool inflationary pressures) that was expected to stop at 71, but actually stopped at 69, can hardly be considered a policy planning failure. On the other hand, an unplanned decline in aggregate economic activity from an index of 100 to 71 must be considered a bad mark against the economic policy makers. Furthermore, different policies are clearly required to combat a mild decline to an index of 69, and to combat a severe decline, say, to 30. That is, from the point of view of policy, the differences between a slowdown and a recession may not be significant, whereas the differences between two "recessions" can be very serious.

Following this approach, Ilse Mintz concluded that if a business cycle

2. *Washington Post*, April 9, 1972.

recession is defined to include retardations in growth, then the leading index shows a one-to-one match at every "cyclical" peak and trough since 1946;[3] it leads at nearly every turning point and does not lag at any; and it gives no false leading signals. Thus, while a decline in the leading index has always signaled a weakening of the economy, these signals sometimes have been followed only by retardations in growth.

Third, I do not find Hymans' rules for testing the forecasting powers of the CLI useful. He fails to take advantage of major efforts to use various smoothing devices to improve the forecasting powers of CLI—principally the technique I have referred to as monthly cyclical dominance, MCD. If he used such a smoothing technique, he would get a much different scorecard from the one he derives with his direct use of seasonally adjusted data and the unrealistic set of rules that he applies to testing those data.

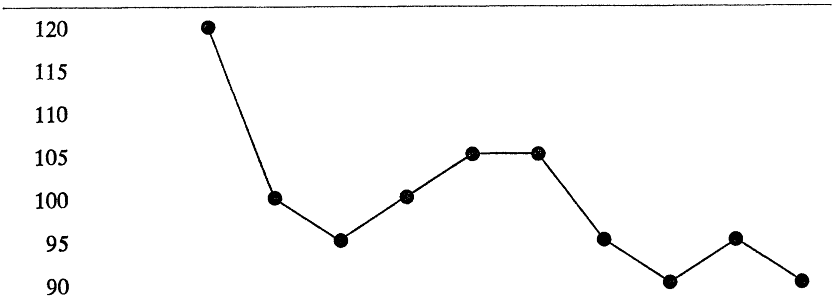
These rules can produce results that are quite contrary to common sense. The four hypothetical examples in Figure 1 illustrate how their application can (a) label a true peak as false (panel A), (b) label a false peak as

3. Ilse Mintz, "Dating American Growth Cycles," in Victor Zarnowitz (ed.), *The Business Cycle Today*, Fiftieth Anniversary Colloquium I (Columbia University Press for the National Bureau of Economic Research, 1972), pp. 39–88.

Figure 1. Hypothetical Examples of Errors under Hymans' Rules for Detecting Cyclical Turning Points from the Composite Leading Indicator
NC = no change; P = peak coming; T = trough coming

A. True Peak Labeled as False

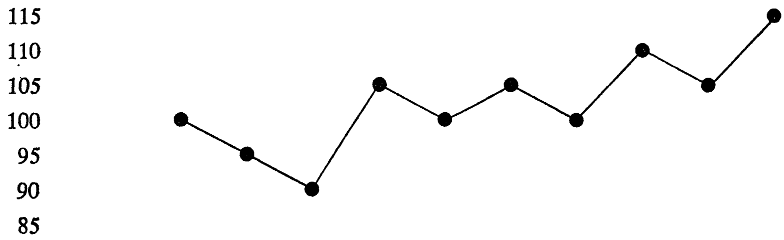
Month	*t−9*	*t−8*	*t−7*	*t−6*	*t−5*	*t−4*	*t−3*	*t−2*	*t−1*	*t*
Prediction	...	*NC*	*P*	*P*	*NC*	*NC*	*NC*	*P*	*P*	*P*
Value	*120*	*100*	*95*	*100*	*105*	*105*	*95*	*90*	*95*	*90*



Hymans' rule: At $t-7$ a peak would be predicted, but this prediction would be canceled at $t-5$ and given again at $t-2$. This process yields a false peak signal and a lead of 4 (assuming the business cycle peak occurs at t).

B. False Peak Labeled as True

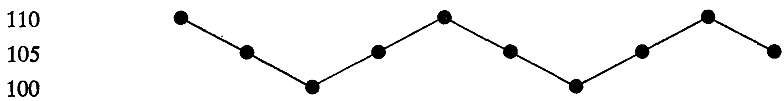
Month	*t−9*	*t−8*	*t−7*	*t−6*	*t−5*	*t−4*	*t−3*	*t−2*	*t−1*	*t*
Prediction	...	*NC*	*P*	*P*	*P*	*P*	*P*	*P*	*P*	*P*
Value	*100*	*95*	*90*	*105*	*100*	*105*	*100*	*110*	*105*	*115*



Hymans' rule: Since two consecutive rises do not follow $t - 7$, the peak prediction made at $t - 7$ will not be revised.

C. Successive False Peaks

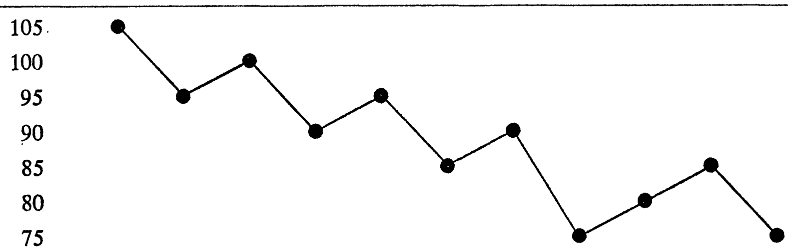
Month	*t−9*	*t−8*	*t−7*	*t−6*	*t−5*	*t−4*	*t−3*	*t−2*	*t−1*	*t*
Prediction	...	*NC*	*P*	*P*	*NC*	*NC*	*P*	*P*	*NC*	*NC*
Value	*110*	*105*	*100*	*105*	*110*	*105*	*100*	*105*	*110*	*105*



Hymans' rule: At $t - 7$ a peak prediction would be made. It would be canceled at $t - 5$ only to be issued again at $t - 3$. The result is two false peak predictions in a very short time.

D. The Missed Peak

Month	*t−9*	*t−8*	*t−7*	*t−6*	*t−5*	*t−4*	*t−3*	*t−2*	*t−1*	*t*	*t+1*
Prediction	...	*NC*	*NC*	*NC*	*NC*	*NC*	*NC*	*NC*	*NC*	*T*	*T*
Value	*105*	*95*	*100*	*90*	*95*	*85*	*90*	*75*	*80*	*85*	*75*



Hymans' rule: At no time would a peak be predicted under the Hymans rule since in no case are there two successive downward changes.

Source: Discussant's examples based on text discussion.

true (panel B), (c) reduce the actual lead time (panel A), (d) generate false peaks (panel C), and (e) completely miss turning points (panel D).

Table 5 of Hymans' paper clearly shows that a lot of spurious turning points were generated by his reliance on these rules. It contains eighty-two predicted and seventy-six realized turning points over a period of two hundred eighty-five months—an average of less than four months between any two successive points. During the same period there were ten NBER turning points and perhaps six growth cycle turning points—a total of sixteen, with an average of nearly eighteen months between any two successive points. Tables 16 and 18 indicate that even the SLI and alternative SLI have averages of a little over seven and four months, respectively. It would appear that Hymans' system is designed to forecast something different from the NBER turning points, even after growth cycle turning points are added.

Hymans seems to understand this. In discussing his spectral analysis approach, he writes, "The shorter periodicities in the range of two to eight months can hardly be interpreted as components of the business cycles that are the focus of this paper." Yet his turning point test appears to be designed to test cycles of these shorter periodicities.

I have asked John Early, a BLS staff member who has done some work on spectral analysis, to comment on some technical aspects of Hymans' SLI series.

John Early: Hymans' use of spectral analysis to construct a spectral leading indicator index (SLI) is both interesting and useful. However, the method raises some questions about the stability of the results. With under three hundred months of data, there are very few repetitions of cycles as long as sixty months in duration. But, as the following data show, more than half of the weight of the index is carried by components with less than eight observed cycles. Thus, the results could be sharply influenced by the inclusion of additional data.

Periodicity (months)	*Weight*	*Observations*
60	0.22	5
40	0.33	7.5
30	0.27	10
24	0.18	12.5

A second problem arises in evaluating the forecasting performance of the

SLI and the composite leading indicator. The predictions are examined within the sample period used to construct the series. Thus the choice of weights and leads are affected by the very cycles that are being predicted. It might be desirable to examine the forecasting accuracy using the best estimates of the weights and leads available prior to the specific turning point that was being questioned. The problem of within-sample bias seems to be less severe for the CLI since it does not make use of lags, which, as already mentioned, may be quite unstable, and since this use of lags is the primary source of improved performance by the SLI.

General Discussion

Thomas Sargent noted that the leading indicators approach to forecasting imposes restrictions that contradict the theory of optimum prediction. In particular, neither a variable nor a function of that variable can be a leading indicator for itself. Yet such functions have been shown to be useful predictors. Sargent also offered an example of a simple case of distributed lag relationships in which a variable, *X*, is exogenous and should be used to predict another variable, *Y*, yet where spectral techniques would show *Y* to be the leading indicator. He noted that when phase statistics vary across frequency, as Hymans reported, it is a symptom of the case in which variables are connected by distributed lags. He suggested that no inference about whether one series should be used to help predict another can be drawn merely from inspecting the phase statistic at a few frequencies.

Thomas Juster pursued Shiskin's concerns about the number of false signals in the indicators reported by Hymans. He agreed that Hymans' filter rules were inappropriate and that the magnitude of a change in the indicators should be taken into account as well as their direction. If rules were devised that tied identification of a change in the phase of a cycle to the size of the movement in the indicators, changes would be called for far less frequently than Hymans reports and the number of false signals would be sharply reduced.

Hymans replied that while additional filtering, such as Juster recommended, would remove some false signals from the record, it would also cut down the forecasting lead of the CLI. He thought that the major difference between the CLI and SLI was that the SLI had adequately filtered the data to start with. He agreed with Juster that experimenting with quantitative rules for interpreting the indicators would be an interesting follow-up to the present work. Hymans also pointed out that his Table 5 should not be read as predicting a cyclical turn on average every five months. Rather, the table shows the number of months when good forecasting would have signaled peak predictions relative to the number when it would have predicted no change in the cyclical phase.

Arthur Okun and Hymans discussed the problem of a correct prediction that is later reversed. Okun noted that in such a case, Hymans was grading the correctness of the first signal by what was forecast subsequently rather than by what actually happened. Hymans replied that consistency forced one to reject such confusing signals, regardless of whether they were correct originally.

Okun and Shiskin then discussed identifying and determining the severity of recessions. Okun suggested that leading indicators should ideally be sensitive enough to distinguish important downturns from retardations or pauses like those in 1956 or 1967. Shiskin replied that a good way of sorting the significant cases from the others would be to ask how policy would be affected. In that sense, he argued that there would be little difference between the 1967 mini-recession and the recessions of 1961 and 1969–70. However, Okun noted that the unemployment rate as an indicator would show an important difference between 1967 and the other two periods, which would have had significant implications for policy.

Charles Holt called attention to a methodological inconsistency in using spectral analysis, which assumes a linear dynamic system, to analyze turning point characteristics of the economy, which are thought to be nonlinear dynamic phenomena. One can approximate any nonlinear system with a linear system, but the application of spectral analysis should be made with this inconsistency in mind. He noted that Shiskin's observation that lead times for indicators are apparently different at peaks and at troughs points up an example of the nonlinearity in the economy's cyclical behavior.