Volume Title: Indicators of Business Expansions and Contractions

Volume Author/Editor: Geoffrey H. Moore and Julius Shiskin

Volume Publisher: NBER

Volume ISBN: 0-87014-444-8

Volume URL: http://www.nber.org/books/moor67-2

Publication Date: 1967

Chapter Title: An Explicit Scoring System for Business Cycle Indicators

Chapter Author: Geoffrey H. Moore, Julius Shiskin

Chapter URL: http://www.nber.org/chapters/c1268

Chapter pages in book: (p. 8 - 28)

II

An Explicit Scoring System for Business Cycle Indicators

1. CRITERIA APPROPRIATE FOR SELECTING INDICATORS

In this review an attempt has been made to develop and apply a reasonably complete and explicit scoring system to aid in the selection and classification of indicators. Such a scoring system can be helpful in systematizing and testing professional judgments in selecting indicators, especially when the judgments are made by different investigators or by the same investigator at different times. It is also helpful in appraising the performance of different series in different types of situations. For example, has the behavior of a given indicator been similar or different at revivals as compared with recessions? Are the indicators which usually have long leads as reliable in other respects as those which usually have short leads? An explicit scoring system can also be used to help select series for various types of composite indexes or other special purposes. Finally, and perhaps most important, the information about each series that is provided by a scoring plan can help the analyst interpret the series' current performance in the light of its past behavior.

The objective of a scoring system for business cycle indicators is limited to evaluating their performance in relation to business cycles, and especially their usefulness in short-term forecasting. It is not concerned with other uses of the series, such as in studies of long-term growth or in governmental administration.

In their 1938 report on business cycle indicators, Mitchell and Burns specified the following characteristics for an ideal indicator of cyclical revivals and recessions.

1. It would cover half a century or longer, thus showing its relation to business cycles under a variety of conditions.
2. It would lead the month around which cyclical revival centers by an invariable interval—say three months, or better, six months. It would also lead the central month of every cyclical recession by an invariable interval, which might differ from the lead at revival.
3. It would show no erratic movements; that is, it would sweep smoothly up from each cyclical trough to the next cyclical peak and then sweep smoothly down to the next trough, so that every change in its direction would herald the coming of a revival or recession in general business.
4. The cyclical movements would be pronounced enough to be readily recognized, and give some indication of the relative amplitude of the coming change.
5. It would be so related to general business activity as to establish as much confidence as the nature of such things allows that its future behavior in regard to business cycles will be like its past behavior.[1]

Mitchell and Burns also noted the importance of having up-to-date figures, good seasonal adjustments, and detailed records of the indicator, its components, and related series.

These criteria were followed in selecting the 1938, 1950, and 1960 lists of indicators, various quantitative measures being devised to implement their application. In the present study we have divided the criteria into six broad types: (1) economic significance in relation to business cycles, (2) statistical adequacy, (3) conformity to historical business cycles, (4) consistency of timing during business cycles, (5) smoothness, and (6) currency.

Each type appears to have an essential role in a system for judging indicators. Economic significance implies that the behavior of a

[1] *Business Cycle Indicators*, Vol. I, pp. 165–166.

particular activity is both well understood and important in the theory of business cycles, i.e., an indicator's performance has a rational explanation. This provides some assurance that it will perform in the future about as well as in the past. On the other hand, unless the historical record of conformity and timing supports this theoretical role, the indicator's claim to a high rating is rendered doubtful; indeed, one might question any key role assigned to it in a business cycle theory.

Statistical adequacy ensures that a series will continue to measure the economic process it is intended to represent equally well during future business cycle fluctuations. In this sense it is an adjunct to the theoretical and historical requirements.

In current business cycle analysis a smooth series is more useful than an irregular one, because in the former a change in direction is more likely to denote the beginning of a new cyclical phase. In an irregular series a new trend must generally run for several months before one can be assured that a new cyclical phase has begun. Irregular series can be smoothed by various statistical devices, but these often impart biases of one sort or another and usually involve a loss of currency. In this sense smoothness and currency are related.

A series must, of course, be available promptly if it is to serve as a useful current indicator. A series which met all the other criteria well, but did not become available until five or six months after the period covered, is apt to be of little use in deciding whether a business cycle turn is imminent or under way, in determining the particular sectors in the economy in which weakness lies, or in selecting the appropriate countercyclical actions to take.

The conversion of these general types of criteria to an explicit scoring system involves many thorny problems. How, for example, can economic significance be judged? It cannot be defined here in terms of relative importance in gross national product, because some large sectors of GNP, such as the service industries, do not have a proportionate role in generating or contributing to cyclical fluctuations, and many relevant aspects of economic activity, such as the functioning of the credit markets, are not recorded in the national product accounts. A more appropriate definition of economic significance in this context would be the role in the cyclical process of the particular activity represented by the series. But in the absence of general agreement on a theory of business cycles, or a working model that reproduces their essential features, is it really possible to discriminate among series on this basis? Again, with most statistical producers each making a case for the reliability of their series, how can the statistical adequacy of different indicators be scored? In racking up a score for historical performance, what weight should be given to the record of conformity to past business cycles as compared with the consistency of past leads or lags? How much should a series be penalized because the most recent month's data are not available when current economic conditions and prospects are being reviewed?

There is certain to be an arbitrary element in the answers given to such questions, especially when the answers are put in quantitative terms. To a large extent, the results will reflect the judgment of the analyst, so that in the end an explicit scoring system may only shift some of the judgmental elements from the final stage of the selection of indicators to an earlier stage, where these implications are less clear. Yet an explicit scoring system does have important merits. It forces the investigator to specify what are his judgments of various properties of each indicator, and it provides a better basis for other investigators to review his work, extend its application, and improve upon it. The detailed results of applying a scoring plan should be valuable in aiding one's understanding of the theory, historical behavior, and methodology underlying the final set of indicators. By pointing to specific deficiencies, the scoring process may, indeed, promote the further development

and improvement of our economic intelligence system.

It may be of interest to note that each of our six criteria has a bearing on the selection of data to be used in an econometric forecasting model. Economic significance surely embraces the idea that the variable is appropriate to include in the equations of the model. Measures of conformity to business cycles help to identify variables that, say, have more to do with long-run growth than with short-run fluctuations in the economy. Lagged variables are a necessary complement to a forecasting model; hence consistent leads or lags are features to be sought. Statistical adequacy, smoothness, and prompt availability at the time the model is to generate a forecast are naturally consequential to the model builder. This is not to say, of course, that the properties we have attempted to weigh in selecting indicators would be weighed in the same way in constructing an econometric model. The results might be quite different. But the criteria are broadly relevant to both approaches, and it is difficult to think of any criterion that is relevant to one but not to the other.

The method of this investigation has been to convert the general criteria listed above into an explicit scoring system with weights assigned by the authors. A separate scoring plan is set up for each criterion, under which a perfect indicator would earn 100 points. The six scores are averaged to obtain a single composite score.

It is recognized that this approach cannot provide an automatic or mechanical method of selecting indicators. The indicators we have selected are not simply those with the top scores. We believe the scheme does provide some objective tests of the quality of different indicators from the standpoint of their value in forecasting, and hence it exercises some control upon the judgment of the investigators in making the final selections. It also puts into clearer perspective the characteristic behavior and limitations of each indicator, and this knowledge is of potential value in the forecasting process itself when the movements of a given indicator are being evaluated. Finally, it may suggest ways in which these or other indicators can be improved to make them more useful for the present purpose.

2. ECONOMIC PROCESSES OR INDIVIDUAL INDICATORS?

A question that arose at the beginning of this investigation was whether to score individual series or groups of series representing broadly defined economic processes. Handling groups of closely related series as a unit has the advantage of avoiding the uncertain and arbitrary elements characteristic of individual series, such as the period they cover or the effects of a few extreme values on measures of cyclical behavior. In the 1950 NBER study, measures of the behavior of individual series were obtained first, then closely related groups of series were evaluated, and finally individual indicators were selected to represent each group. Among other things, this procedure made it possible to select series that had only recently been constructed and had only a brief record, if they were clearly superior in coverage or other respects to closely related series that had a longer record.

In terms of our list of six criteria, a case can be made for a combination of both approaches. That is to say, broadly defined economic processes might be evaluated under some of the criteria and individual indicators under others.

Under the criterion of economic significance it would seem best to evaluate groups of series representing a general type of activity, because theories which purport to explain business cycle phenomena do not ordinarily refer to particular indicators, but rather to generalized economic processes. On the other hand, since methods of compilation vary from series to series and change from time to time, statistical adequacy, smoothness, and currency pertain more directly to individual series. The value of

a particular series for analyzing short-term business trends and prospects depends importantly on these properties.

The problem of the timing and conformity criteria is more complex. Here there is some advantage in scoring groups, because of the light that one series throws upon the behavior of another. For example, the recorded average lead at business cycle peaks of the unemployment rate (inverted) for married males is 15 months. But, because this series begins only in 1954, the average is based on observations at only two peaks (one a lead of 19, the other of 11 months). Related series that are available for a much longer period, such as the total unemployment rate, suggest that the tendency to lead is genuine but that the typical lead is closer to 4 months than to 15. Hence the evidence supporting the conclusion that unemployment, even of married males alone, tends to rise prior to a business cycle peak is much stronger than the limited evidence supplied by that series itself. Nevertheless, the historical record of each series is difficult to disregard. Can the same weight be given to a series with a poor conformity record as to another with a good conformity record merely because they fall in the same economic group?

Another troublesome question is how to group series for this purpose. The 1950 indicator study used a classification designed by Mitchell for his work on business cycles.[2] This classification reflected Mitchell's extensive knowledge of business cycle theory and history, the distinctions he had observed in the cyclical behavior of different processes, and the statistical series that were then available to fill the classifications. Today, the advances in our knowledge, the changes in cyclical behavior that have occurred, and the new statistical series that have been constructed all point to the need for a new classification.

On balance, the problems involved in combining group scoring with individual series scoring appeared too great to make it worthwhile, and each series has been scored independently. However, in making a final selection of indicators it is clearly of great importance to take into account the evidence provided by closely related series, and we have attempted to follow this practice.

For this, as well as for other purposes, we have devised a simple economic process classification. We believe that the following nine types would be generally recognized as strategic processes in business cycles, although different economists would divide them differently, substitute or add some other items, and certainly weigh them differently in their thinking:

1. Employment and unemployment
2. Production, income, consumption, and trade
3. Fixed capital investment
4. Inventories and inventory investment
5. Prices, costs, and profits
6. Money and credit
7. Foreign trade and payments
8. Federal government activities
9. Economic activity in other countries

Many series in the first two groups are measures of aggregate economic activity and are used to describe the broad movements of the business cycle and to determine the dates when business expansions and contractions begin or end. These two and the next four groups also include factors which are credited by students of the business cycle with a causal role in the cyclical process, that is, in the cumulative processes of expansion and contraction and in the reversal from expansion to contraction and from contraction to expansion. The last three groups represent processes which are not generally considered responsible for cyclical fluctuations in the United States, but which nevertheless importantly affect their pattern, amplitude, and duration. Many of the groups, in particular groups 6 and 8, include factors that reflect the implementation of governmental policy with respect to recession, unemployment, inflation, or other features of business cycles.

[2] See *Business Cycle Indicators*, Vol. I, pp. 214–215.

3. ECONOMIC SIGNIFICANCE

Economic significance is an essential element in selecting—as well as in using—business cycle indicators. It is the *sine qua non*. No matter how excellent an indicator's historical performance or statistical basis, it cannot be given great weight in analyzing short-term economic developments unless it measures or represents an activity with a key role in the cyclical process. But economic significance in this sense is also a most difficult element to evaluate in quantitative, objective terms. We have tried and rejected several alternative plans, including an attempt to consider each series' role in various explanations of the business cycle. We have also considered the possibility of omitting this element entirely from the explicit scoring scheme. In the final analysis, all the series on our list are economically significant; otherwise they would not have been considered at all. The question is whether it is better to try to recognize different degrees of significance, rough as the results are apt to be, than to leave this element out of the scoring plan and implicitly equate all series in this respect. Our judgment is that explicit scores for economic significance will make a contribution to the scoring system, enhancing both its rationalization and its discriminating power. We recognize, however, that there may be wide disagreement on the particular plan used for scoring as well as on the individual scores assigned for this element.

In order to evaluate the economic significance of indicators for business cycle analysis, it seems necessary to take into account at least two factors. One is the role of a given economic process in theories or hypotheses that purport to explain how business cycles come about or how they may be modified or controlled. The other is the breadth of coverage of a particular series representing that process.

The eight (or nine) types of economic process listed above appear to include all the variables deemed significant in modern business cycle analysis. Some factors that in earlier times attracted attention, such as variations in the weather or in the frequency of sunspots, are not covered. But the list does provide for "real" as well as "monetary" factors, for consumption as well as investment, for inventory as well as fixed capital investment, for costs as well as prices, for resource utilization rates as well as profits, for governmental as well as private activity, for international as well as domestic developments.

The eight categories are, however, rather broad. We can get closer to a specification of economic variables that are represented by statistical series if we subdivide them. Table 2 presents such a subclassification (column 1), developed with an eye to the types of economic indicators that are available and have proven to be of analytical interest in business cycle studies. The subclassifications are not exhaustive; others might well be added should occasion warrant. Moreover, they are not mutually exclusive. Some series might be classified in more than one of the subcategories. Thus, prices of industrial materials represent costs as well as prices, and so do interest rates. Orders for durable goods pertain to investment in capital equipment, to consumer purchases of autos and other durables, and to the accumulation of inventories of steel and other materials.

Despite these limitations, the subcategories constitute a list of economic variables that are of strategic interest in business cycle analysis, forecasting, and policymaking. Some of the available economic indicators represent these variables in comprehensive fashion; others are confined to particular sectors, components, or aspects. In columns 2 and 3 of the table the 122 indicators reviewed in this study are classified according to whether, conceptually, they are "broad" or "narrow" in coverage. Those that purport to cover the entire economy, or the "cyclically sensitive" portion of it, or major fractions thereof (such as all corporate activity, or total consumption or invest-

TABLE 2
Classification of 122 Series According to Economic Process and Breadth of Coverage

Type of Economic Process (1)	Broad Series[a] (2)	Narrow Series[b] (3)
1. Employment and unemployment		
Marginal employment adjustments		Av. workweek, mfg.
	Nonagri. placements	Accession rate, mfg.
		New hires, mfg.
		Rehires, mfg.
	Temporary layoffs	Layoff rate, mfg.
	Initial claims, unempl. insur.	
Job vacancies	Nonagri. job openings	
	Help-wanted ads	Executive help-wanted ads
Employment	Nonagri. man-hours, employees	
	Nonagri. man-hours, with job	
	Nonagri. man-hours, at work	
	Nonagri. empl., estab. survey	
	Nonagri. empl., household survey	
	Nonagri. empl., commodities	Nonagri. empl., services[c]
Unemployment	Unempl. rate, total	Unempl. rate, under 5 wks.
		Unempl. rate, 5–14 wks.
		Unempl. rate, 15+ wks.
	Unempl. rate, insured	
	Unempl. rate, married males	
2. Production, income, consumption, and trade		
Production	GNP, current $, expend. est.	
	GNP, constant $, expend. est.	
	GNP, current $, income est.	
	GNP, constant $, income est.	
	Industrial production	Steel ingot production
		Pass. car production
Income	Personal income	Labor income in mining, mfg., and construction
Consumption and trade	Bank debits outside N.Y.C.	
	Final sales, current $	
	Final sales, constant $	
	Mfg. and trade sales	Mfrs.' sales
		Wholesalers' sales
		Truck tonnage hauled
	Retail sales	Cons. expend., dur. goods
3. Fixed capital investment		
Formation of business enterprises	Net business formation	
	New incorporations	
New investment commitments	New orders, dur. goods	New orders, mach. and equip.
		New orders, mach. tools
	Constr. contracts, total	Housing starts
		Building permits, housing
	Contracts and orders, plant and equipment	New cap. approp., mfg.
		Comm. and indus. contracts
Backlog of investment commitments		Unfilled orders, dur. goods
		Cap. approp. backlog, mfg.

TABLE 2 (*Continued*)

Type of Economic Process (1)	Broad Series[a] (2)	Narrow Series[b] (3)
Investment expenditures	Gross priv. dom. invest., total	
	Gross priv. dom. invest., bus. sec.	
	New pl. and equip. expend.	
	Mach. and equip. sales and bus. constr. expend.	Prod. dur. equip. sales
		Equipment production
		New constr. expend., bus.
4. Inventories and inventory investment		
Inventories	Mfg. and trade inventories	Mfrs.' inventories, total
		Mfrs.' inventories, fin. goods
Inventory investment and purchasing	Change, bus. inventories	
	Change, mfg. and trade invent.	Purch. mat., % reptg. higher inventories
		Change, purch. mat. inventories
		Buying policy, prod. mat.
		Vendor performance
		Change, unfilled orders, dur.
		Change, dept. store stocks on hand and on order
5. Prices, costs, and profits		
Sensitive commodity price indexes		Indus. mat. prices
Stock price indexes	Stock prices, 500 common stocks	
Wholesale price indexes	Wholesale prices, exc. farm products and foods	Wholesale prices, mfd. goods
Retail price indexes	Consumer price index	
Unit labor costs	Labor cost per $ of real corp. GNP	Labor cost per unit, mfg.
Profits and profit margins	Corp. profits, total	
	Profits to income orig., corp.	Profits to sales, mfg.
		Price to labor cost, mfg.
6. Money and credit		
Flows of money and credit	Change, money supply and time deposits	
	Change, money supply	
	Total private borrowing	Change, consumer instal. debt
		Change, bank loans to bus.
		Change, mortgage debt
		New nonfarm mortgages recorded
	Corp. gross savings	Stock offerings, mfg. corp.
		Stock sales, N.Y.S.E.
Outstanding debt		Cons. instal. debt outst.
		Comm. and indus. loans outst.
Bank reserves	Free reserves	
Money market interest rates		Treasury bill rate
		Corp. bond yields
		Treasury bond yields
		Municipal bond yields
Interest rates on business loans and mortgages		Bank rates on bus. loans
		Mortgage yields, residential
Credit difficulties	Liab. of bus. failures	No. of large bus. failures
		Delinq. rate, all instal. loans
		Delinq. rate, direct auto loans

TABLE 2 (*Concluded*)

Type of Economic Process (1)	Broad Series[a] (2)	Narrow Series[b] (3)
7. Foreign trade and payments	Balance of payments	
	Merchandise trade balance	
	Exports	Export orders, dur. goods
		Export orders, machinery
	Imports	
8. Federal government activities	Surplus or def., income and prod.	
	Cash surplus or deficit	
	Cash receipts	
	Cash payments	Defense purchases
		Defense oblig., total
		Defense oblig., procurement
		New orders, defense products
		Military contracts

Note: This classification does not take into account the *statistical* coverage of the series; i.e., what fraction of the population it purports to cover is actually reported. For full titles of series, see Appendix G.

[a] Economy-wide; nonagricultural; manufacturing and trade; total corporate; commodity, consumption, or investment aggregates.

[b] Manufacturing; other sectors or components narrower than those listed under note a.

[c] See text.

ment), are considered "broad."[3] Series that pertain to a single industry (e.g., manufacturing) or to minor components of the "broad" series are placed in the "narrow" group.[4]

One of the theoretical advantages of broad coverage is that it provides protection against substantial changes in cyclical behavior that may arise from such factors as technological developments, changing consumer tastes, or the rapid growth or decline of single products or industries. A broad economic indicator may continue to perform well, or at least in representative fashion, even though some of its components deteriorate in this respect. Thus, one would expect total retail sales to continue to be an important indicator, while department store sales alone might diminish in significance.

The table tells us something about the economic significance of the several indicators and their relationship to one another in terms of economic coverage. We have not attempted to distinguish degrees of significance among the economic categories in which the series are classified. Hence we haye given all series in the "broad" column the same score, namely 75. Similarly, all series in the "narrow" column are given a score of 50. The assignment of the same score to all series within a column means, of course, that some obvious differences in coverage are ignored.

The result of this process is that all of the indicators considered in this review get scores of either 75 or 50 for economic significance, depending upon their coverage.[5] This does not mean that any economic series whatever would be entitled to such a score. The indicators

[3] This does not imply that their coverage is broad from a statistical point of view. Some of the "broad" series are based on limited statistical samples. This aspect of coverage is considered in section 4, below.

[4] The classification of service industry employment poses a problem, since it is currently a larger component of nonagricultural employment than the "cyclically sensitive" commodity-producing employment. Since the commodity sector could, while the service sector could not, be considered representative of employment from the cyclical point of view, the service series is relegated to the "narrow" group.

[5] The levels 75 and 50, while arbitrary, yield an average score for economic significance roughly similar to that for the other five categories in the scoring plan.

under review are already a highly selected group. Within it, it did not seem feasible or especially useful to make fine distinctions with respect to their significance for cyclical analysis and forecasting, apart from their breadth of coverage.

4. STATISTICAL ADEQUACY

A sound statistical method of compilation is a necessary condition for a good indicator, since it provides some assurance that the figures can be relied upon, in the future as in the past. Some of the considerations involved apply chiefly to the series as currently issued; others to the historical data.

An important requirement is that the series be based upon a reporting system. The aggregate it reflects should be obtained by summing reports on their activities made by respondents. Thus retail sales should be obtained by summing reports from retailers or consumers, manufacturers' orders by summing reports from manufacturers or buyers, employment by summing reports from employers or members of the labor force. This requirement may seem obvious, but some important series, such as the index of industrial production, or the index of net business formation, or even gross national product, are based largely upon indirect sources.

Good coverage is a second requirement. With careful collection, editing, and processing of returns, complete coverage is obviously best. But complete and accurate coverage is often difficult, partly because of its cost and partly because of the unwillingness of some respondents to report. When a sample is necessary, it should be a probability sample, for then a measure of the error that arises in estimating the universe from partial coverage can be provided. Other kinds of samples may also give good results, but they have the disadvantage that their accuracy cannot be expressed in quantitative terms. The mere size of the sample is not a sufficient criterion, for it is sometimes possible to obtain a more accurate series from a smaller sample than from a larger one. When reporting is poor and processing costs large, more careful handling of a smaller number of returns may result in a reduction of the reporting and processing errors that is greater than the error introduced by the smaller sample.

A statistic should cover the full period it represents; for example, a series representing a monthly total should cover the full month. For reasons of economy, some series refer to only one week or even one day of the month. This is a different kind of sample and, like other kinds, leaves something to be desired, especially when the figures are not accompanied by a measure of the error resulting from this short cut.

It is customary to release current statistics before all the returns are in. These "preliminary" figures are later revised at various intervals when more complete information becomes available. Thus, a substantial proportion of the returns in the sample of manufacturers from whom sales and orders data are collected become available by the twentieth of the month following that covered by the data, but the sample is more complete by the end of the month, and some returns come in during the following month. More nearly complete estimates for all manufacturing are available in the *Annual Survey of Manufactures*, while full detail for various products may become available only once every five years in a census of manufactures. The accumulation of better information leads to revisions.

Methods of seasonal adjustment also lead to revisions. For example, the moving average methods commonly used require more years of data to estimate seasonal factors for the current year than are available at the time the estimates for the current year are made. When the additional data do become available, the seasonal factors and the seasonally adjusted data are revised. It is clear that a

measure of the extent and significance of seasonal and other revisions is a desirable supplement to a statistical series.[6]

A measure of the total error, from all sources, to which each figure is subject is therefore an important requirement for a sound statistical series. For business cycle studies, errors of the estimates of change are frequently more important than errors of the estimates of level. Apart from sampling errors, respondents make errors in reporting, statistical agencies make errors in processing or estimating, and compositors make errors in printing. Partly for theoretical reasons and partly because of costs, it is difficult to derive a measure of the total error. None of the statistical series covered in our review now provides such a measure, though efforts to obtain them in the survey of unemployment and related series are under way. The existence of closely related series often provides important clues to the errors to which any one of them is subject.

The factors discussed above apply to the currently issued data. But to appraise the usefulness of a series as a cyclical indicator, a historical record must also be available. From this point of view, two considerations are relevant: the period for which the series is available and the comparability of the series over time. A series going back many years obviously provides more information on its cyclical behavior than a short series. The producers of statistical series are often confronted with the dilemma of improving a series or maintaining its historical comparability. Both are important. Improvement is often necessary to reflect the forces which the series is designed to measure, to expand coverage, or to catch up with changes in the structure of the economy. However, breaks in comparability reduce the value of the series for historical analysis.

Finally, a full account of the survey methods (content of survey form, collection procedure, sampling, editing, coding, and other processing), the coverage both in terms of the respondents and the period, and the seasonal and related adjustment methods is an essential requirement for statistical series. On the other hand, although seasonal, trading-day, and smoothing adjustments contribute to the usefulness of statistical series in their role as business cycle indicators, electronic computers have made such auxiliary data available for all the business indicators considered in this review. For this reason, it seems unnecessary to score these factors.

5. CONFORMITY

A principal factor in judging the historical performance of indicators is their conformity to past business cycles. The National Bureau's index of conformity provides a simple measure of how faithfully each series has followed the business cycle chronology. A series that has

[6] When are preliminary statistics sufficiently accurate to be useful in interpreting short-term business trends and prospects? One approach to this question is to compare the magnitude of revisions with the magnitude of changes in the underlying trend. Thus, preliminary estimates of month-to-month changes might be considered acceptable if the subsequent revisions are on average smaller than the average month-to-month change in the underlying trend. The following formula expresses this criterion:

$$\frac{\bar{R}}{\bar{C}} < 1,$$

where $\bar{R}$ is the average revision of the month-to-month changes and $\bar{C}$ is the average month-to-month change in the trend-cycle component of the final series, both averages being taken without regard to sign.

Other ways of appraising the usefulness of preliminary figures are to determine whether revisions affect the direction of the month-to-month changes in the MCD curves or whether they substantially alter the proportion of "significant" month-to-month changes. For the derivation of $\bar{C}$ and the MCD curves, see Julius Shiskin, *Electronic Computers and Business Indicators*, Occasional Paper 57, New York, NBER, 1957.

An extensive study of revisions of gross national product and its components is being carried out by Rosanne Cole as part of the National Bureau's study of short-term forecasting.

risen during every business expansion and declined during every contraction will have an index of +100; a series which has declined during every expansion and risen during every contraction will have an index of −100. Systematic leads and lags are taken into account. For example, if a series typically leads, an allowance is made for the average (median) lead in computing the conformity index. The conformity index is a type of correlation coefficient between the cyclical fluctuations in each series and those in aggregate economic activity. The business cycle chronology used in this study is recorded in Appendix F.

The length of the record upon which the conformity index is based, i.e., its statistical significance, needs to be taken into account. Following a scheme used in the 1950 NBER study, we have computed the probability, for a series covering a specified number of cycles, that the conformity index would reach the observed level. This probability, rather than the conformity index itself, is used in the scoring plan.

The conformity index, however, misses several important aspects of conformity. First, it is unaffected by extra cycles. For example, during the business cycle expansion of 1949–53, the extra declines in many series during 1951 are disregarded in computing conformity indexes, since only the change between levels at business cycle troughs and at business cycle peaks are considered. In analyses of current business cycle developments such extra movements may be misleading, as when a current decline in a series is interpreted as a signal of recession but eventually proves to be an extra cyclical phase having no general significance. Second, the conformity index does not distinguish early lapses in conformity from recent lapses. Yet recent lapses are a matter of greater concern than those which occurred many years ago, because of the ever-changing forces that shape the business cycle. Third, the conformity index does not take the amplitude of the cycle into account. A series which reveals a cycle clearly and decisively is more useful, other things equal, than one whose cyclical movements are mild and difficult to distinguish from other types of fluctuation.

Hence in our scoring plan the conformity index and its probability are supplemented by measures of extra cycles, recent lapses, and amplitude. It would be desirable, also, to take account of the relation between the amplitude of the movements in an indicator, particularly in the early months of a recession or recovery, and the amplitude of the business cycle, but that has not been done in the present study.[7]

6. TIMING

One of the most firmly established findings of business cycle research is that the cyclical movements in many different economic activities typically occur at somewhat different times; that is, some series lead and others lag. These relationships are of vital significance to the forecaster. Their nature and stability must therefore be recognized in a scoring system.

Several important aspects of timing are considered: the consistency with which timing comparisons of the same type (leads, rough coincidences, or lags) occur at successive business cycle turns;[8] the variability in length of (say) the lead, as measured by the dispersion about the average; the recency of defections from characteristic behavior; the presence of long-run shifts in timing; and the difference between timing at business cycle peaks and at troughs.

[7] Measures of this type were in a 1958 NBER report, *Measuring Recessions* (reprinted in *Business Cycle Indicators*, Vol. I, Chap. 5, pp. 120–161).

[8] Leads and lags are measured (in months) by comparing the date of the cyclical peak (or trough) in the series with the monthly business cycle peak (or trough) date, according to rules established for matching these turns. Leads (lags) are timing comparisons in which the cyclical turn in the series precedes (follows) the business cycle turn by one month or more. Rough coincidences overlap these, being defined as leads or lags of three months or less, including exact coincidences.

As in the case of conformity, a probability scheme for judging consistency of timing, developed in the 1950 study, is used here. It involves calculating the probability that as large a proportion of leads (or rough coincidences or lags) as that observed during the business cycle turns covered by the series could occur by chance. This allows for the fact that a given proportion of leads is more significant the larger the number of cycles covered. It takes account also of the possibility that the series may skip a particular cycle and hence fail to produce any timing comparison. Errors of the opposite kind, i.e., *extra* cycles, are allowed for under conformity, as noted above.

The length of the average lead or lag is of obvious importance in using indicators. It is difficult, however, to designate an ideal lead time. It might be argued, for example, that a lead of thirty months is too long to be useful in planning policies to combat inflation or recession, while a lead of three months does not provide enough time to muster the forces necessary for an effective attack. However, much depends on the variability about the average and the availability of other information. No one indicator should be considered in isolation, but rather as one among many factors in a developing situation. Let us suppose that one indicator leads by twenty-four months, another by twelve, and a third by six. Then early-warning signals of the end of expansion (contraction) could be observed as they unfold, and countercyclical policies systematically prepared. Furthermore, a series that lags regularly can also be helpful. It may confirm the doubtful signals of other series, or it may reveal developments that will eventually set the stage for a reversal of the cyclical tides.

In the light of these considerations, it does not seem possible to prefer one average lead time to another,[9] and we have, therefore, not assigned scores on the basis of these averages. However, the average (median) lead or lag is used to classify the indicators. That is, subject to qualifications noted below, an indicator is classed as leading if the median of all the timing comparisons is a lead of two months or more *and* if the number of leads, considered in relation to the number of business cycle turns covered by the series, is sufficient to reach an "acceptable" level on the probability scale mentioned above. Similarly, a series is classed as lagging if its median timing is a lag of 2 months or more and the number of lags is significantly large. A series is roughly coincident if it fails to satisfy the above criteria but exhibits a significant number of rough coincidences (timing observations that lie within a range of ± 3 months). Hence most roughly coincident series have median timing of -1, 0, or $+1$ months, though medians of ± 2 are possible (i.e., if with a median of ± 2 the series fails the probability test for leads or lags but qualifies for rough coincidences). The same significance level is used for all three classifications.[10]

---

[9] There is one rather exceptional situation that affords a basis for preference. In a few series with very long average leads, some of the individual leads approximate or even exceed the length of the corresponding business cycle phase. For example, the rate of change in the money supply reached its trough in December 1959, some fourteen months before the business cycle trough of February 1961, with which it is matched. Here the upturn in money even preceded the downturn in business (May 1960), and the series rose throughout the contraction. Such leads may, if they occur frequently, raise the question whether the series is positively or inversely related to business cycles. In any case, they make for difficulties in interpretation. For this reason, an unusually long average lead may be disadvantageous, other things equal.

[10] This is the same significance test used in the 1950 study, which was based upon a tabulation of the leads and lags of some 404 series with "acceptable" conformity. Cf. *Business Cycle Indicators*, Vol. I, p. 209. The record covered business cycles between 1854 and 1938. A similarly comprehensive record that includes the period since 1938 has not been compiled, but a tabulation limited to a group of 90 selected indicators for the period through 1961 gave results quite similar to the earlier one, considering the greater selectivity of the sample.

7. CURRENCY AND SMOOTHNESS

Prompt availability is an essential requirement of an indicator. In the counsels of practical men, a series that is current commands attention, whereas one that is out of date is apt to be disregarded. Thus, if a figure on corporate profits, surely a dynamic factor in business cycles, does not appear promptly after the quarter covered, it may not even be posted on a score sheet of current developments. On the other hand, more promptly available indicators, such as retail sales and unemployment, will get heavy weight in the decision-making process. It is unfortunately the case that judgments of the current outlook are often based upon a poor representation of series because they must be made before reports for all the indicators are in.

The availability of daily figures (prices of industrial raw materials or of common stocks) or weekly figures (retail sales or initial claims for unemployment insurance) must be counted in favor of an indicator. Such data are often helpful in making an early estimate for the current month, and occasionally they are useful directly in determining a cyclical turning point. As a rule, the shorter the time unit in which data are reported, the shorter the lag in their availability to the analyst.

A smooth series has the advantage that it is more likely than an irregular one to give prompt notice of the beginning of a new cyclical phase. An ideal indicator would be one that changed direction only in the event of a recession or recovery. This explains in large part the efforts of economists and statisticians to disentangle the underlying cyclical movements of economic series from other types of change, especially seasonal and irregular fluctuations.

Economic series vary a great deal in their smoothness. None meets the ideal, though some come close to it. Indeed, many indicators that meet other criteria well are highly irregular, for example, housing starts or the liabilities of business failures. A study of 150 economic series showed that the month-to-month changes are on average "cyclically significant" in only about 25 per cent; even in these, not all of the month-to-month changes are cyclically significant.[11] Since irregularities dominate the short-term changes of the other series, comparisons over longer spans must usually be made to detect cyclical changes.

This points up the fact that higher degrees of smoothness can be achieved either by observing the series over longer intervals, e.g., comparing the latest month's figure with the figure six months ago, or by consolidating the figures into longer time units, as for example by a moving average. Either way there is a loss in timeliness, because recent changes in direction are obscured. Because of this relationship between smoothness and currency, both factors must be taken into account in evaluating indicators. For example, there is generally a net advantage in weekly over monthly data that are equivalent in other respects, even though weekly data are often much more erratic in their movements. The reason is that the weekly data can, if necessary, be converted to four- or five-week moving totals, in which case they will be as smooth as the corresponding monthly data, and as a rule more up to date. The advantage is even greater, of course, when the monthly data pertain to only one week in the month. Monthly data possess similar advantages over quarterly figures.

It should be noted that preliminary estimates for most series are usually more erratic than the final estimates. Revisions of seasonal

[11] "Cyclical significance" is defined as a condition in which the average change in the irregular component ($\bar{I}$) of a series is smaller than in the cyclical component ($\bar{C}$). The ratio, $\bar{I}/\bar{C}$, generally diminishes as the span over which change is measured increases. See *Business Cycle Indicators*, Vol. I, p. 607.

adjustments generally yield smoother adjusted data, and so do revisions that result from a larger reporting sample. Our measures of smoothness are based upon the historical record, not the current record. Eventually, we may be able to devise measures appropriate to current data. A justification for the present method is that for comparisons among series, historical smoothness is probably highly correlated with current smoothness.

Measures of smoothness, currency, and timing should, ideally, be considered in relation to one another. The loss of currency that occurs when an irregular series is smoothed may be offset, in some degree, by a long lead. Similarly, a series that is not promptly available but has a long lead may be as useful as a series that comes out promptly but has a shorter lead. For example, the gross accession rate for manufacturing is not available as promptly each month as the broader series, nonagricultural placements, and its irregular fluctuations are larger, but the accession rate nearly always has shown a longer lead. If the accession rate were smoothed by, say, a three-month moving average, its loss of currency on this account as well as its publication lag would still not quite offset the advantage of its longer lead.[12] This illustrates the desirability of evaluating the leads and lags of indicators after smoothing them to an equivalent degree and taking into account both their publication lag and the additional lag produced by smoothing. We have not been able to do this systematically in this study, but Appendix E shows the effect on the median lead or lag of the 72 leading, coincident, and lagging indicators of adjusting them to a roughly equivalent degree of smoothness by simple moving averages. In general, the effect on the sequences among the indicators is not great, though the leads of some highly erratic series are substantially reduced.

8. COVERAGE OF SERIES UNDER REVIEW

The present review has covered 122 series. Included are the 80 U.S. indicators presently published in *Business Cycle Developments* and some 42 additional series which have at one time or another been considered for inclusion. The additional series cover a wide range, but there is some emphasis on areas where the present list is relatively thin, such as financial indicators.

The "population" of series upon which the review is based, therefore, is by no means a random collection of economic data. This was equally true of the more extensive reviews made previously by the National Bureau. In all cases the series considered for review represented processes that were judged, in some degree, economically significant in the analysis of business cycles. In some instances, however, processes that might have been considered significant were omitted, or poorly represented, because adequate monthly or quarterly data are not available. The possibility remains that some important indicators have been overlooked. Suggestions to this effect would be welcomed by the authors.

[12] At the nine business cycle turns covered by both series, 1945–61, the accession rate led placements on all but one occasion (when they were coincident) and the median lead was 4 months. The $\bar{I}/\bar{C}$ ratios for the two series are, respectively, 3.20 and 1.23. Smoothing by a three-month moving average would reduce the former to about the level of the latter, namely, 1.11. Allowing for a one-month lag on account of the centering of the three-month average and another month for publication lag would cut the median effective lead of the accession rate relative to placements to 2 months, and it would have led placements at 5 turns, coincided twice, and lagged twice. Placements are analogous to the new-hires component of the accession rate, and this component often moves later than the accession rate—i.e., later than the other component, rehires. For a discussion of some of these relationships, see the report by Charlotte Boschan in the 44th Annual Report of the National Bureau, New York, 1964, pp. 106–110.

9. APPLICATION OF THE SCORING PLAN ILLUSTRATED

The way this scoring method was applied can be illustrated by the series, new orders received by durable goods manufacturers, an important leading indicator. It received the relatively high score of 78.

The scores assigned for each element considered in evaluating a series were posted on a special "scoresheet." Many of the entries were derived from detailed statistical records, posted on a "worksheet for timing, conformity, and smoothness." Most of these, in turn, were based upon the business cycle measures developed by the National Bureau of Economic Research and described in the volume *Measuring Business Cycles* by Burns and Mitchell. The scoresheet and worksheet for new orders of durable goods are shown here as Tables 3 and 4. This record covers the full period for which the series is available, 1920–65. For series that start before 1948, such as this one, a second scoresheet covering the period 1948 to 1965 is also prepared (not illustrated here), so that the scores for all series can be compared for a common, recent period, 1948–65. (See Appendix D.) The detailed instructions for scoring are given in Appendix A.

First, the series was assigned 75 for economic significance, the higher of the two levels adopted for this category. The strategic role of investment in capital goods in business cycles is generally recognized. This series records an early stage of the investment process, not only in the sense that it records orders of equipment largely for future delivery, but also in that it includes orders placed for materials to be used in the production of capital goods. The materials, such as steel and lumber, enter into structures as well as equipment, and the finished goods, such as automobiles, aircraft, and machinery, include those purchased by business enterprises, consumers, and government. Hence, directly and indirectly, it has very broad coverage, warranting its assignment to the "broad" rather than the "narrow" group.

With respect to the next item, statistical adequacy, the series is assigned the full 20 points for "reporting system" since it is compiled from reports sent to manufacturing companies for statistical purposes only. It accumulates another 15 points because it is based upon a probability sample, and 10 more because the data cover the full month, and not some portion of it.

The compiling agency has not yet published a measure of the magnitude of revisions or the sampling error for this series; hence, no points are assigned for these items. On the other hand, full credit (5 points) is given for the availability of a detailed description of the series.

In order to give some weight to the length of the historical period covered by a series, 2 points are assigned for each five years covered, up to 50 years. The earliest segment of the series on new orders for durables begins in 1920; hence it is counted as covering 45 years, and scores 18 points.

With respect to comparability over time, it is necessary to take account of the serious breaks in the series in 1928 and 1938. There is some doubt, also, about its comparability before and after 1953, because of improvements in compilation procedures, but we decided to consider it essentially continuous since 1948. The series therefore scores 4 points for comparability. For statistical adequacy as a whole, the series achieves a mark of 72 points ($20 + 15 + 10 + 5 + 18 + 4$).

The timing, conformity, and smoothness worksheet (Table 4) shows that new orders conformed to the business cycle in all the phases for which data are available—nine expansions and nine contractions. That is, it rose during every expansion and fell during every contraction, when allowance is made for the

TABLE 3
Sample Scoresheet for Indicators

SERIES: Manufacturers' New Orders, Durable Goods Industries,
NBER, 1920-28; NICB, 1929-38; OBE, 1939-46; Census, 1947-65

BCD # 6 NBER # 6,84 & 6,91

			SCORE
I	ECONOMIC SIGNIFICANCE (20%) (importance in business cycle theory, broadness of coverage in the economy)		75
II	STATISTICAL ADEQUACY (20%)		
	Current data:		
	1. Reporting system (based upon reports obtained directly from respondents, or administrative records such as tax reports, or similar sources)	20	
	2. Coverage of process (universe, probability sample or other types of sample)	15	
	3. Coverage of time unit (full calendar month or quarter; or selected week or day)	10	
	4. Revisions (measure of frequency and magnitude)	0	
	5. Measure of error (availability of, for total error or sampling error)	0	
	6. Description (completeness of, for compilation or estimation methods)	5	
	Historical data:		
	7. Duration (time period covered)	18	
	8. Comparability of historical series		
	Absence of breaks in comparability	0	
	Last segment 15 yrs. or longer	4	
			72
III	CONFORMITY (20%) (to the business cycle)		
	1. Conformity probability (that as large a number of movements in the series which conform to cyclical movements in general business could occur by chance)	60	
	2. Extra turns (absence of cyclical turns in the series that do not correspond with turns in general business)	12	
	3. Recent lapses (absence of lapses in conformity since 1948)	10	
	4. Amplitude (extent to which series undergoes clear cyclical movements)	6	
			88

TABLE 3 (*Concluded*)

		SCORE
IV TIMING (20%) (consistency with which timing comparisons of the same type--leads, rough coincidences, or lags--occur at successive business cycle turns)		
Peaks		
1. Timing class	Leads	
2. Probability (that as large a number of leads--or rough coincidences or lags--as that observed during the business cycle turns covered by the series could occur by chance)	55	
3. Dispersion (variability in the lengths of lead or lag)	0	
4. Recent lapses (recency of defections from typical behavior)	20	
		75
Troughs		
1. Timing class	Leads	
2. Probability	58[a]	
3. Dispersion	14	
4. Recent lapses	20[a]	
		92
Peaks and Troughs		84
V CURRENCY (10%)		
1. Promptness (interval after period covered that statistic becomes available)	80	
2. Daily or weekly reports (availability of daily, weekly, or 10-day reports)	0	80
VI SMOOTHNESS (10%) (estimate of the likelihood that a change in direction denotes the beginning of a new cyclical phase)		60
		AVERAGE SCORE
VII SUMMARY		
Peaks } based on leads at peaks and leads at troughs		76
Troughs } based on leads at peaks and leads at troughs		79
Peaks and Troughs } based on leads at peaks and leads at troughs		78

[a] Timing accepted for rough coincidences: probability score 7, lapses 0.

Note: Percentage figures in parentheses next to major headings show weights assigned in computing average scores. Scoresheet actually used did not include explanatory statements.

TABLE 4

Sample Worksheet for Measures of Timing, Conformity, and Smoothness

Manufacturers' New Orders, Durable Goods Industries
NBER, 1920-28; NICB, 1929-38; OBE, 1939-46; Census, 1947-65

BCD # 6.
NBER # 6,84 & 6,91

Business Cycle Peak	June 1869	Oct. 1873	Mar. 1882	Mar. 1887	July 1890	Jan. 1893	Dec. 1895	June 1899	Sep. 1902	May 1907	Jan. 1910	Jan. 1913	Aug. 1918	Jan. 1920	May 1923	Oct. 1926	Aug. 1929	May 1937	Feb. 1945	Nov. 1948	July 1953	July 1957	May 1960
Lead(-)or Lag(+)at Peak(mos.)[1]															-4	-11		-5	-35	-3	-6	-19	-13
Expansion Conformity[2]															+	+	+	+	+	+	+	+	+
Business Cycle Trough	Dec. 1870	Mar. 1879	May 1885	Apr. 1888	May 1891	June 1894	June 1897	Dec. 1900	Aug. 1904	June 1908	Jan. 1912	Dec. 1914	Mar. 1919	July 1921	July 1924	Nov. 1927	Mar. 1933	June 1938	Oct. 1945	Oct. 1949	Aug. 1954	Apr. 1958	Feb. 1961
Lead(-)or Lag(+)at Trough(mos.)[1]														-6	-2		0	-2	-2	-4[c]	-11	-3	-1
Contraction Conformity[2]															-	-	-	-	-	-	-	-	-

	No. of Business Cycle Turns Covered	No. of Extra Specific Cycle Turns	Total No. of Comparisons Made	Number of: Leads: 4 Mos. or More	Leads: 1-3 Mos.	Coincidences: Exact	Coincidences: Rough[3]	Lags: 1-3 Mos.	Lags: 4 Mos. or More	Per Cent of Business Cycle Turns Covered: Leads	Coincidences: Exact	Coincidences: Rough[3]	Lags	Median Lead(-) or Lag(+) (months)	Mean Lead(-) or Lag(+) (months)	Standard Deviation of Leads & Lags (months)	Longest: Lead	Longest: Lag	Timing Probability	Timing Rejected or Accepted for
Peaks	10	1	8	7	1	0	1	0	0	80	0	10	0	-8 1/2	-12.0	10.1	-35	-3	.0202	Leads
Troughs	10	1	9	3	5	1	6	0	0	80	10	60	0	-2	-3.4	3.1	-11	0	.0079 / .2207	Leads / R.C.
Peaks & Troughs	20	2	17	10	6	1	7	0	0	80	5	35	0	-4	-7.5	8.4	-35	0	.00034	Leads

Conformity	No. of Phases Covered	Index	Conformity Probability	Accepted or Rejected
Exp.	9	+100	.0020	Accepted
Con.	9	+100	.0020	Accepted
Full Cycle	18	+100	.0000038	Accepted

Specific Cycle Amplitude per Month	
Rise	+3.14
Fall	-2.45
Rise & Fall	2.80

Period Covered	$\overline{CI}$	$\overline{I}$	$\overline{C}$	$\overline{I}/\overline{C}$	MCD	$\overline{I}/\overline{C}$ for MCD Span	Average Duration of Run: CI	I	C	MCD
Jan.'53-Sept.'65	3.76	3.33	1.51	2.20	3	.66	1.81	1.58	8.44	4.41

1
Timing comparisons are footnoted as follows if they apply:

a
Equal to or longer than the corresponding business cycle phase.
b
Comparison crosses an opposite specific cycle turn.
c
Comparison rejects a competing specific cycle turn..

2
A plus (+) signifies a rise and a minus (-) a decline during a business cycle expansion or contraction, taking into account the median lead or lag of the series.

3
Includes leads of 1-3 months, exact coincidences, and lags of 1-3 months.

Notes: For definitions of timing, conformity, and amplitude measures, see Arthur F. Burns and Wesley C. Mitchell, Measuring Business Cycles, National Bureau of Economic Research, New York, 1949, Chapter 5. Note, however, that the conformity measures used here, which take leads or lags into account, are based on the median lead or lag of the series (in months) rather than the stages formerly used. For descriptions of the probability measures, see Geoffrey H. Moore, Statistical Indicators of Cyclical Revivals and Recessions, Occasional Paper 31, NBER, New York, 1950. For definitions of smoothness (MCD and $\overline{I}/\overline{C}$) and related measures, see Julius Shiskin, Electronic Computers and Business Indicators, Occasional Paper 57, NBER, New York, 1957. The last two references are reprinted in Business Cycle Indicators, Chapters 7 and 17, respectively.

average (median) lead of the series with respect to business cycles. The probability that such a result could have been obtained for a series that is randomly related to the business cycle is calculated to be .002 for expansions and contractions separately, and is still smaller for expansions and contractions combined. Our formula assigns the series the maximum number of points (60) for this measure of conformity. Since there were no lapses in conformity after 1948, an additional 10 points is scored here. However, the series had an extra cyclical turning point at peaks (January 1951) and an extra turning point at troughs (January 1952); hence it receives only 12 of the 20 possible points on this item. The average specific cycle amplitude per month is a moderately high 2.80 per cent, which yields 6 points out of a possible 10. Thus, for the various items grouped under conformity, new orders for durables total to 88 points ($60 + 12 + 10 + 6$).

Under timing, calculations are made for peaks and troughs separately and combined, on the basis of the entries on the worksheet. At peaks, this series led in all cases except 1929, where no timing comparison could be made; at troughs, leads also prevailed except for the coincident timing in 1933 and the absence of a timing comparison in 1927. The median timing was a lead of 8 months at peaks and 2 months at troughs. This record qualified the series as a leader at peaks and at troughs, with timing probability scores of 55 and 58 points respectively out of 60. Because of the high dispersion at peaks (standard deviation, 10.1 months), it failed to receive any points for this item. But, with a smaller dispersion at troughs (standard deviation, 3.1), it received 14 out of 20 possible points. Twenty additional points were credited because a timing comparison was made at every business cycle turn since 1948. Thus new orders for durables received 75 points for timing at peaks, 92 for timing at troughs, or an average of 84 for peaks and troughs combined.

This series is credited with 80 points for currency since an advance release comes out before the twentieth of each month, with data for the preceding month. There is no weekly estimate; therefore, no points could be added for this.

Several measures of smoothness are given on the worksheet.[13] The one used for scoring monthly series is the number of "months for cyclical dominance" (MCD), which for new orders is 3. Hence, our rules yield 60 points for smoothness.

Our rules assign equal weight to economic significance (75), statistical adequacy (72), conformity (88), and timing (84), and half weight to currency (80) and smoothness (60). The average for new orders for durables is therefore 78, which is the final score.

It must be emphasized that the absolute value of the scores for each criterion or for all combined has no significance. Its function is to aid in the relative evaluation of one series compared with another. In judging the scores for the individual series, therefore, it is important to consider their place in the frequency distribution of the scores for all the series covered by the study, and their relation to the average scores for the various items. These are given in Table 5.

The frequency distributions show that no series received a final score as high as 90 and that only 4 scored 80 or better. The average score was 62. Fifty-seven received a score of 75 for economic significance. This is consistent with the fact that the sample of series included in this study was selected with a view to their broad economic significance. According to the scheme for judging the statistical adequacy of the series, it turned out that only 4 had a score of 80 or over and none over 90.

Twenty-two series received a score of 90 or more for conformity, but 15 series scored below 50. The record on this criterion is again related to the design of the sample, which was

[13] For a discussion of the meaning and uses of these measures, see *Business Cycle Indicators*, Vol. I, pp. 535–545, 604–609.

TABLE 5
Distribution of Scores for 120 Indicators

Score	Economic Significance	Statistical Adequacy	Conformity	Timing	Smoothness	Currency	Final Score
	(number of series)						
100	...	...	...	...	26	17	...
95–99	...	...	6	...	...	...	...
90–94	...	...	16	2	...	...	...
85–89	...	1	15	3	...	...	1
80–84	...	3	14	11	32	44	3
75–79	57	17	17	5	...	...	7
70–74	...	25	18	5	...	...	14
65–69	...	23	7	11	...	...	33
60–64	...	29	3	6	27	1	20
55–59	...	12	2	12	...	...	15
50–54	63	5	7	9	...	10	8
45–49	...	3	2	4	...	...	10
40–44	...	1	2	12	13	34	5
35–39	...	...	...	5	...	...	4
30–34	...	1	1	2	...	...	...
25–29	...	...	1	6	...	7	...
20–24	...	...	4	5	7	...	...
15–19	...	...	1	5	...	...	...
10–14	...	...	...	6	...	...	...
5–9	...	...	1	9	...	...	...
0–4	...	...	3	2	15	7	...
Total	120	120	120	120	120	120	120
Average score	62	66	72	48	62	61	62

Note: Based on scores for full period covered by series. Two series (export orders of durables and export orders of machinery) are omitted because complete set of measures is not available.

selected mainly with an eye to series that conform regularly to business cycles, but it included some strategic series that conform poorly.

The timing scores are not as high as the conformity scores. Only 2 series had a timing score of 90 or better for peaks and troughs combined, and nearly half had timing scores under 50. The reason for the wide distribution is the same as for conformity.

For currency, or promptness of publication, only 17 series received a perfect score. These are series which are available by the twentieth of the month following that covered by the data and for which weekly or daily figures are available. Twenty-six series met our highest standard for smoothness.

Of the various elements scored separately, the highest average was attained by conformity (72) and the lowest by timing (48). Since no adjustment for such differences was made before computing the average score for each series, the elements with the highest average scores implicitly received more weight. It is to be noted, however, that if conformity is averaged with timing to attain a single score for historical performance during business cycles, the differences among the factors is small: all average between 60 and 66.[14]

[14] As a partial test of the scoring plan, a cumulated random series (i.e., a series with random first differences) was constructed from a table of random numbers to provide monthly observations over a 45-year period. This artificial series was adjusted so that

10. INTERPRETATION OF SCORES

The assigned scores should be considered rough rather than precise measures of the relative usefulness of different series in analyzing short-term business conditions and prospects. Despite our attempt to provide an objective appraisal of economic time series, many arbitrary elements, where judgments could differ considerably, enter into the design and execution of the plan. This is particularly true of the assignment of weights for the various factors, such as economic significance or statistical adequacy. If other investigators prepared a similar scoring program, it would no doubt differ in many respects from ours. We would venture the guess, however, that there would be a fairly high correlation between their scores and ours.

The precise value of the final score, therefore, has limited significance. It is a convenient symbol of success, like the grades assigned to students on an examination. The results depend to a degree upon the questions asked, the points assigned to each, and the judgment of the examiner in evaluating the responses. But the scoring plan conveys other information, not revealed in the final score. The scheme sets forth the many different factors that need to be taken into account in appraising indicators. Altogether, 20 different items are rated for each series. While there can be disagreement about their relative importance, we believe there would be general agreement on their relevance. Furthermore, the scores assigned point to particular merits and limitations of series, and in this way may be of assistance to both producers and users of statistics. A detailed examination of them may suggest which series are worth improving and in what respects improvements are needed. Thus, series which have low ratings for promptness of publication, but high ratings for economic significance and historical record, might be worth speeding up. Series whose principal defect is a large erratic movement might be intensively studied to find the causes and possible cures for this defect. In any case, users of the statistics can be forewarned of their limitations, and qualify their interpretations accordingly.

---

the expected average month-to-month change without regard to sign was equal to 5 per cent. Its MCD proved to be 3, and its smoothness score was 60. Scores for conformity and timing were obtained on the assumption that the series began at four alternative hypothetical dates (January 1919, June 1919, January 1920, and June 1920). In all four cases, the conformity score was 0. In three cases the timing score was also 0, and in one it was 26. This indicates, as we would expect, that series with cyclical properties but basically unrelated to the U.S. business cycle are unlikely to achieve scores that approach those achieved by the economic indicators included in this study.