On the Dating of Business Cycles

Victor Zarnowitz

Journal of Business, Volume 36, Issue 2 (Apr., 1963), 179-199.

Your use of the JSTOR database indicates your acceptance of JSTOR's Terms and Conditions of Use. A copy of JSTOR's Terms and Conditions of Use is available at http://uk.jstor.org/about/terms.html, by contacting JSTOR at jstor@mimas.ac.uk, or by calling JSTOR at 0161 275 7919 or (FAX) 0161 275 6040. No part of a JSTOR transmission may be copied, downloaded, stored, further transmitted, transferred, distributed, altered, or otherwise used, in any form or by any means, except: (1) one stored electronic and one paper copy of any article solely for your personal, non-commercial use, or (2) with prior written permission of JSTOR and the publisher of the article or other text.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Journal of Business is published by University of Chicago Press. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://uk.jstor.org/journals/ucpress.html.

Journal of Business
©1963 University of Chicago Press

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor@mimas.ac.uk.

©2001 JSTOR

http://uk.jstor.org/
Tue Sep 11 13:54:03 2001

ON THE DATING OF BUSINESS CYCLES*

VICTOR ZARNOWITZ†

METHODS and results of economic research that have empirical content require testing against the relevant "facts," and the more important they are (for analysis, prediction, or policy) the greater the need for the tests. In principle, testing is a never ending business: any sufficient input of new data of the proper sort enables one to challenge afresh this approach or that proposition. The dating of business cycles is certainly important enough to warrant a great deal of attention. An article by George W. Cloos in the previous issue of this *Journal* poses in its title the question "How Good Are the National Bureau's Reference Dates?" and starts out by contending that these dates "have not been tested adequately." In a broad sense, as as these introductory remarks suggest, this contention is incontrovertible. But it is sometimes well to be reminded of a truism. Cloos's paper, by what it does and fails to do, illustrates the need for serious attempts to review the business-cycle chronology critically.

Cloos proposes the following alternatives to the NBER postwar chronology: (*a*) the cyclical timing of GNP; (*b*) the cyclical timing of the FRB index of industrial production; (*c*) a set of turning points based primarily on four comprehensive monthly series on income, employment, output, and sales. I shall discuss each of these approaches and compare them with the National Bureau dating method. First, however, it will be necessary to give some thought to the substance of the cycle-dating problem and the rationale of the NBER analysis. Following the comparison of the chronologies, I shall address myself to the main points raised by Cloos and then to some of his sundry criticisms of the Bureau's method and measures. The final section will present an illustration of the problems encountered in reference dating, based on materials pertaining to the 1954 trough.

THE RATIONALE OF THE NATIONAL BUREAU METHOD

Cloos's presentation of the NBER dating method abstracts completely from the basic view of the business cycle that underlies the Bureau's measurements. Indeed, he fails to make explicit the rationale not only of the Bureau's but also of his own approach. However, it is not difficult to recognize that the concept of "aggregate economic activity" implicit in Cloos's argument is quite different from the NBER concept. The latter is very comprehensive, the former much more restrictive.

The definition of business cycles used in the National Bureau studies as a working concept to be tested by observation refers to a type of fluctuation "occurring at about the same time in many eco-

* This paper is in large part a reply to George W. Cloos, "How Good Are the National Bureau's Reference Dates?" *Journal of Business*, XXXVI (January, 1963), 14–32. I am grateful to Arthur F. Burns, Geoffrey H. Moore, and Alexander Pitts, Jr., of the National Bureau of Economic Research for providing illustrative data and supporting information, as well as helpful comments.

† Associate professor of finance, Graduate School of Business, University of Chicago, and member of research staff, National Bureau of Economic Research.

nomic activities."[1] To examine the validity of this concept, it is necessary to inspect the behavior of time series portraying a great variety of economic process—not only in the sphere of production of "final" goods and services or employment or income or sales but also in such areas as money and banking, prices, security markets, business population changes, and others. If business cycles "are due to the predominance of agreement in timing among specific cycles," then "the peaks of positive and troughs of inverted specific cycles in a representative collection of time series must occur in clusters, and so also must the troughs of positive and the peaks of inverted specific cycles."[2] The NBER reference chronology originated in the effort to identify such definite "clusters" of turning points as could be accepted to mark the culmination of a business-cycle expansion or contraction. The Bureau's studies showed that peaks and troughs of series representing a broad array of economic processes are indeed heavily concentrated in certain periods and around particular dates. The latter tend to indicate the months in which aggregate economic activity reached its peak or through levels.[3]

The Bureau's "definition" is thus deliberately open, implying that not enough is yet known about the subject to exclude from study any large group of economic processes. This position is not "antitheoretical." It is shared, for example, by such a theorist as Hicks who says early in his book on the trade cycle that "it is a mistake to begin one's investigation with a definition of the kind of fluctuation which one is going to regard as basic—deciding whether one is going to regard the cycle as being fundamentally a fluctuation in employment, or output, or prices, or interest rates, or money supplies. It is better to allow the definition to emerge as the theory develops."[4]

Cloos, on the contrary, decides that it is better to define the business cycle at the outset as the cycle in GNP or even the FRB index which covers only the production of factories, mines, and gas and electric utilities, and now embraces roughly 35 per cent of GNP.

It is true that considerations of simplicity and definiteness argue in favor of using a single comprehensive series or a combination of a few such series to measure aggregate economic activity and its cyclical timing. Viewed as a complex of interdependent partial processes, total economic activity represents a multidimensional macrocosmos, the levels and changes of which are not directly measurable. The concept of general business activity was no doubt always more closely related to comprehensive indexes of production or output values than to measures of other aspects of the econ-

[1] Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (New York: National Bureau of Economic Research, 1946), p. 3.

[2] Wesley C. Mitchell, *What Happens during Business Cycles: A Progress Report* (New York: National Bureau of Economic Research, 1951), p. 10.

[3] If the collection of the series used is really representative, these dates will agree with the evidence of the turning points in the most comprehensive and important aggregates available, for they are basically reflecting the central tendency in the timing of these aggregates and their principal components. However, a reference date thus derived need not coincide exactly with the turning point in any particular aggregate that some might want to single out (e.g., industrial production). The dates express the consensus of a large group of series and usually lie in a central position with respect to the array of their specific turns. They are analogous to an average and the leads and lags measured from them are analogous to deviations from the average. This is a convenient property as far as the use of reference dates for cyclical timing comparisons is concerned.

[4] J. R. Hicks, *A Contribution to the Theory of the Trade Cycle* (Oxford, 1950), p. 2.

omy; and this was not altered by any theoretical or empirical results of business-cycle studies. Hence, if one wishes to approximate the concept by a single measure, the gross national product (perhaps preferably in constant prices) appears as the most logical choice.

However, weighty reasons remain that still support a decision to determine the timing of general business revivals and recessions—the "reference dates"—by analyzing the behavior not of the GNP alone but of a large group of selected comprehensive series. (To be sure, the GNP figures, where available, are to be included in this group and given much weight in deciding upon the dates.) To arrive at a business-cycle chronology in which one can have as much confidence as the nature of the problem allows, it is not desirable to rely on the evidence of a single series, even if it is as comprehensive and important as GNP. The use of GNP figures alone does not necessarily assure the selection of a unique set of reference dates, as will be illustrated in the following section; and it definitely means that less information is employed than the amount conveyed by a group of properly chosen series.

It is only by a comprehensive analysis of economic time series that the crucial characteristics of business cycles, such as their scope, can be ascertained. Before determining the date of, say, a business-cycle peak, one has to decide whether a contraction of business-cycle proportions has occurred or is about to occur, which usually requires looking at many series representing diverse economic activities. This aspect of the dating problem is a basic one, but it is often overlooked.

If sole reliance on GNP (or any other single indicator) is not desirable, neither is it necessary for technical reasons. Complementary methods of summarizing the behavior of diverse economic processes are available. Alternating clusters of specific peaks and troughs in a sufficiently large and representative group of series typically permit identification of months in which the turns of either type are most highly concentrated. Diffusion indexes (usually calculated as the percentage of the given set of series that are rising) provide a means for describing the collective behavior of a group of series that cannot otherwise be meaningfully aggregated, and for measuring the degree of consensus among their expansions and contractions. In using these techniques for the purpose of dating the business-cycle turns it is of course still necessary to take careful account of differences in the economic significance of the series. To do so may be difficult, but it is certainly possible.

THE TIMING OF GNP

In determining the reference dates, the Bureau relies primarily on monthly series. Burns and Mitchell present a large amount of detailed evidence in support of the Bureau's preference for monthly rather than annual, or indeed quarterly, series for purposes of cyclical analysis. Monthly reference dates are therefore basic, and the NBER procedure is to set them first (with the aid of evidence supplied largely by monthly but also by quarterly series). The quarterly reference dates are then made to match the monthly ones as follows: If the monthly choice falls on the middle of a quarter (say August), that quarter is always taken as the quarterly reference date. If it falls on the first or third month of a quarter, the quarterly turn is placed either in the quarter in which the reference month is located or in the quarter adjacent to that month.[5]

[5] Burns and Mitchell, *op. cit.*, p. 80.

The differences between the reference chronology and GNP timing must be viewed in the context of this technique. Even if the turn in the quarterly GNP coincides with the reference quarter, it may diverge from the monthly reference date by as much as two months, since this is the maximum divergence between the Bureau's monthly and quarterly reference dates (see, e.g., the case of the 1953 peak in Table 1).

ing, given the above technical considerations that may account for them.

The main practical difficulty with the GNP figures, as this analysis suggests, is that they are not available monthly, whereas monthly reference dates are needed. But it must also be noted that even the quarterly figures are, for some sectors, largely interpolated or extrapolated. Such estimates are inevitably subject to uncertainties and therefore

TABLE 1

CYCLICAL TIMING OF GROSS NATIONAL PRODUCT AND INDUSTRIAL PRODUCTION COMPARED WITH NBER MONTHLY AND QUARTERLY REFERENCE DATES, 1948–61

Year of Peak (P) or Trough (T)	NBER Reference Dates		Dates of Cyclical Turns in		Lead (−) or Lag (+), in Months, Relative to NBER Monthly Reference Dates		
	Monthly (1)	Quarterly (2)	Gross National Product (GNP) (Quarterly) (3)	Industrial Production (IP) (Monthly) (4)	NBER Quarterly Reference Dates[a] (5)	GNP[b] (6)	IP (7)
1948 P.....	November	IV	IV	July	0	0	−4
1949 T.....	October	IV	II	October	+1	−5	0
1953 P.....	July	II	II	July	−2	−2	0
1954 T.....	August	III	II	April	0	−3	−4
1957 P.....	July	III	III	February	+1	+1	−5
1958 T.....	April	II	I	April	+1	−2	0
1960 P.....	May	II	II	January	0	0	−4
1961 T.....	February	I	I	January	0	0	−1

[a] Intervals from the midmonth of the reference quarter to the reference month.

[b] Intervals from the midmonth of the GNP peak or trough quarter to the reference peak or trough month.

Source: National Bureau of Economic Research, Inc. The dates of specific cycle turns in GNP and industrial production (IP) are based on currently available data (for GNP: United States Department of Commerce, Office of Business Economics, *U.S. Income and Output*, 1958, and *Survey of Current Business*, July, 1962; for IP: Board of Governors of the Federal Reserve System, *Federal Reserve Bulletin*, October, 1962).

In the postwar years (the period covered by our present quarterly GNP data), only two differences of more than two months are found between the cyclical turns in GNP and the monthly reference dates. These are the leads of GNP of five months and three months at the 1949 and 1954 troughs, respectively (see Table 1). It will be seen that the dating of these two upturns presents special problems that are difficult to resolve. The other divergences are hardly exciting, also to revisions that properly seek improvements in the light of additional information. The revisions sometimes lead to large shifts in timing. Thus the 1949 trough in GNP was recently shifted from the fourth quarter to the second (giving rise to the largest single discrepancy in Table 1; see col. [6]). Later in this paper the frequency and the effects upon timing of the revisions of GNP will be strikingly illustrated for the case of the 1954 trough (see Table 5 below).

GNP can be derived from income as well as from expenditure accounts, the difference between the two being the "statistical discrepancy." Either estimate has about the same claim to accuracy. Yet at the 1954 revival, the income-accounts total reached its low in the first quarter; the expenditure-accounts total, in the second (see Table 4 below). No episodic finding of this sort, of course, should be taken to indicate that GNP figures are to be "rejected" as unreliable; but there is at least a suggestion here that it may not be wise to rely on these data exclusively.[6] The timing of GNP in constant dollars was the same as that of GNP in current dollars throughout the postwar period, but this, of course, need not always be the case. If it is not, a choice would have to be made as to whether to rely on the deflated or the current dollar series.

THE TIMING OF INDUSTRIAL PRODUCTION

The FRB industrial production index led at four reference turns in the postwar period and showed coincident timing at four (Table 1, col. [7]). The 1957 recession shows how poorly one may fare in using a single indicator to select a business turn date, even if the series are as important and generally useful as the FRB index.[7] Output and employment in manufacturing had been undergoing slow declines since late in 1956, as indicated by a study of the FRB production and the BLS employment series both in the aggregate and for a number of major industries. Meanwhile, activity in most other sectors continued to advance. The total industrial production index (including utilities) reached its peak in February, six months before GNP (whose peak falls in the third quarter of the year, centered on August).[8]

Cloos asserts, without citing his evidence, that "there is reason to believe that if a 'monthly GNP' were available, it would move fairly closely with the FRB index."[9] Even if this were granted, what should be done when they diverge (as they certainly may, especially if GNP is in current dollars)? But one may well doubt the validity of the assertion. Indeed, the possibility exists that the differences in behavior between GNP and industrial production may be increasing because of the rising importance of sectors not included in the FRB index, such as services.

Cloos also believes it to be important that the FRB index and the constant-dollar GNP had nearly identical percentage amplitudes in two selected recent periods (1957–60 and 1960–61), although he adds: "Comparisons over time are not always this close."[10] But even on visual inspection it is immediately apparent that the FRB index has consistently larger relative cyclical amplitudes than has the deflated GNP series.[11] This

[6] For an analysis of the difference between the two measures of GNP during the 1953–54 and 1948–49 recessions see *Economic Report of the President* (January, 1955), pp. 84–85.

[7] Incidentally, I am quite sure that the appreciation of the merits of the FRB index, or of GNP, is at least as high among the NBER staff as among any other group of economists; suffice it to note that, after all, not a little work on national income accounts and industrial production statistics was done by men such as Kuznets, Fabricant, Moore, and others. In fact, the Bureau uses these aggregates and their components so intensively that it sounds a bit strange to me when Cloos speaks of their "rejection" by the NBER.

[8] The index underwent several revisions in the postwar period, notably the major ones in 1953 and 1959. The figures used here come from the latest revision completed in 1962. Compare col. (4) in Table 1 with col. (2) in Table 2 for the effects on some of the dates (in 1948 and 1957) of very recent changes in the data.

[9] Cloos, *op. cit.*, p. 23. [10] *Ibid.*

[11] See, e.g., Chart 1B for 1948–62 in United States Department of Commerce, *Business Cycle Developments*, for any recent month.

is hardly surprising, as it is well known that manufacturing is more sensitive cyclically than most other sectors of the economy. However, one should also recall that the FRB index does not cover some industries that are highly cyclical such as freight transportation and construction, and does include the relatively stable utilities.

The FRB index, of course, is not error-free (few economic time series are). It is capable of improvement through revisions to which it is subject from time to time. Also, nearly half of the index is based on man-hours multiplied by monthly interpolated or extrapolated productivity factors. These calculations assume smooth monthly changes in output per man-hour between annual benchmark levels. Such interpolations made by a mathematical formula, which are to a degree arbitrary, can affect the date of a turn in the index. Physical output series are in general more reliable than the adjusted man-hours series from the standpoint of a short-run analysis, and the component of the FRB index made up exclusively of the former deserves to be carefully observed; but its coverage (mainly materials and consumer goods) is considerably narrower than that of the total index.

IS BUSINESS CYCLE DATING POSSIBLE? IS IT NEEDED?

Cloos is apparently not sure whether the selection of reference dates is impossible, possible but wrongly done (by the NBER), or unnecessary. He questions whether the month of a trough can be selected unless also the duration of a separate "revival" phase (which he defines as an "interval between a contraction and an expansion") is identified. He reproves Mitchell and Burns for attempting to do the former, while refusing to do the latter, on the ground that "both troughs and revivals are located by examining clusters of turning points in specific series."[12] But there is no contradiction in the NBER position: the "clusters" for a large group of series typically extend over many months, but they also show, as a rule, definite points of concentration.[13]

The view that there is need for reference dates to determine the order in which different activities join the revival or recession is to Cloos "a puzzler."[14] He denies this need on the ground that the specific turns in the individual series have already been dated. But he fails to note the reason given by Burns and Mitchell, which is that "when the analysis covers hundreds of series, it is clumsy and wasteful to compare the timing of each series with every other; indeed, as clumsy and wasteful as it would be to express the exchange value of each commodity in terms of every other commodity."[15]

What is obviously true is that the sequence of turning points in specific series is *not affected* by a selection or shifting of a reference date. This clearly does not eliminate the practical need for a reference chronology, but it does mean that the consequences of errors and shifts in reference dates for the analysis of timing

[12] Cloos, *op. cit.*, p. 16. The reference here is to excerpts from Wesley C. Mitchell and Arthur F. Burns, *Statistical Indicators of Cyclical Revivals* (National Bureau of Economic Research Bull. 69 [New York: National Bureau of Economic Research, 1938]) (reprinted as chap. vi of *Business Cycle Indicators*, ed. Geoffrey H. Moore [2 vols.; Princeton, N.J.: Princeton University Press, 1961]).

[13] See, e.g., Chart 7.3 in Geoffrey H. Moore, *Statistical Indicators of Cyclical Revivals and Recession* (Occasional Paper No. 31 [New York: National Bureau of Economic Research, 1950]) (reprinted as chap. vii in *Business Cycle Indicators*, *op. cit.*).

[14] Cloos, *op. cit.*, p. 16.

[15] *Measuring Business Cycles*, *op. cit.*, pp. 70–71.

relations among interdependent economic processes are not very troublesome. A common reference point offers a convenient short-cut device for measuring these relations, but the latter are independent of it. This is important because from the standpoint of economic analysis it is the relations among the economic processes that are of principal interest.

Comparisons employing the business-cycle chronology by no means claim the exclusive attention of the Bureau. NBER studies in this field have frequently used comparisons in which specific cycle turns in selected series (e.g., GNP, industrial production, freight traffic) serve as the reference frame. Cloos discovers an example of this in an article by Burns[16] and proclaims that "Burns does not always use the Bureau's reference periods in analysis."[17] What does this prove except that Burns, like others, finds it simpler at times to describe cyclical movements in terms of a specific series? How does it suggest that the "trust that outsiders place in the Bureau's reference quarters does not seem to be shared by its president, Arthur F. Burns"?

AN ALTERNATIVE SET OF MONTHLY REFERENCE DATES

In any event, whether the task is impossible or merely unnecessary, Cloos finally decides to make up his own monthly reference chronology for the postwar period. His dates are based on three series classified by NBER as "roughly coincident"—industrial production, non-agricultural employment, and personal income—plus the series of total business sales substituted for the retail sales used by the Bureau. His method seems very much like that of the Bureau—inspecting selected time series, identifying their turning points, picking the month expressing their "consensus" —except that only a few aggregative series are consulted and no use is made of such tools as the diffusion indexes. The procedure hardly represents an alternative *method* of dating, although it is so labeled.

The remaining five of the nine coincident NBER indicators are excluded on various grounds. "The rate of unemployment," he declares, "is a measure of inactivity rather than activity and is of no use here."[18] This is truly a bold move: Would Cloos ignore the unemployment figures in appraising currently the cyclical situation of the economy? After this, the statement that "Neither prices nor bank debits are direct measures of activity" is anticlimactic. Still, it is a little puzzling. Are these series "indirect" measures of activity? No measures at all? Are we to repudiate all faith in the debits series that, as recounted by George Garvy, used to be referred to by Federal Reserve officials as "the Board's statistics of volume of business!"[19]

For the three series common to the two lists of indicators, Cloos's specific-cycle dates differ from those identified by the Bureau about half the time (cf. cols. [1]-[2], [4]-[5], and [7]-[8] in Table 2). NBER dates lead those of Cloos nine times (by intervals ranging from one to six months) and lag them three times (by three months in each case, all at the 1949 trough).

Some of these discrepancies probably reflect revisions of the data. The NBER dates in Table 2 refer to very recent (end

[16] *Morgan Guaranty Survey*, August, 1961.

[17] Cloos, *op. cit.*, p. 22.

[18] *Ibid.*, p. 22.

[19] George Garvy, *Debits and Clearings Statistics and Their Use* (rev. ed.; Washington, D.C.: Board of Governors of the Federal Reserve System, May, 1959), p. 108.

of February, 1963) versions of the series used, whereas Cloos's dates appear to be based on the data as they have been a year or so ago. NBER dates derived from still earlier materials differ in several instances from the dates listed in Table 2, illustrating shifts in turning points due to data revisions.[20]

Cloos identifies the specific-cycle peaks and troughs in the series after having rounded the figures to three digits. In case of ties he picks the latest month. Table 2 presents results obtained by applying this method to the same data as were used to derive the other NBER dates listed (i.e., to the series as of February, 1963, as noted before.) These results show that the method of rounding the figures frequently yields later dates than those one would pick using unrounded series (at least when applying the NBER criteria of specific-turn selection.) Rounding accounts for half the divergences observed between Cloos and

TABLE 2

SPECIFIC TURNS IN THREE SELECTED AGGREGATIVE SERIES, CLOOS VERSUS NBER DATING, 1948–61[a]

YEAR OF PEAK (P) OR TROUGH (T)	INDUSTRIAL PRODUCTION			NON-AGRICULTURAL EMPLOYMENT			PERSONAL INCOME		
	Cloos (1)	NBER (2)	NBER (Rounded Data)[b] (3)	Cloos (4)	NBER (5)	NBER (Rounded Data)[c] (6)	Cloos (7)	NBER (8)	NBER (Rounded Data)[d] (9)
1948 P....	Oct.	July (−3)	Oct.	Nov.	July (−4)	Sept. (−2)	Oct.	Oct.	Oct.
1949 T....	July[e]	Oct. (+3)	Oct. (+3)	July[e]	Oct. (+3)	Oct. (+3)	July[e]	Oct. (+3)	Oct. (+3)
1953 P....	July	July	July	July	May (−2)	May (−2)	Oct.	Oct.	Oct.
1954 T....	Apr.	Apr.	Apr.	Aug.	Aug.	Aug.	May	Mar. (−2)	May
1957 P....	Aug.	Feb. (−6)	Aug.	Aug.	Mar. (−5)	Aug.	Aug.	Aug.	Aug.
1958 T....	Apr.	Apr.	Apr.	May	Apr. (−1)	May	Feb.	Feb.	Feb.
1960 P....	Jan. or July	Jan. (0 or −6)	Jan. (0 or −6)	Apr.	Apr.	Apr.	Nov.	[f]	[f]
1961 T....	Feb.	Jan. (−1)	Feb.	Feb.	Feb.	Feb.	Feb.	[f]	[f]

[a] Figures in parentheses show leads (−) or lags (+), in months, of the NBER dates relative to Cloos's dates. Where the dates coincide (0), no figures are given. NBER dating is based on series available as of February 28, 1963. These series may have been revised subsequent to Cloos's dating. Cloos's dates are taken from his article in *Journal of Business*, January, 1963, p. 30.

[b] Using index to a whole number instead of one decimal.

[c] Using three digits instead of five.

[d] Using three digits instead of four.

[e] Cloos notes that these series were lower in October, 1949, than in July but believes that this was a temporary dip due to the steel strike.

[f] No specific cycle contraction.

[20] The dates for 1948–58 used in *Business Cycle Indicators*, Vol. I, Appendix B, Table B1, are listed in the tabulation below for all those instances in which they differ from the new NBER dates as given in Table 2, cols. (2), (5), and (8). (Leads and lags are measured in months from the new dates in the usual manner.)

	Industrial Production	Non-agricultural Employment	Personal Income
Peaks:			
1948...	August (+1)	November (+4)	September (−1)
1953...	.....	June (+1)	.....
1957...	March (+1)	July (+3)	.....
Troughs:			
1954...	.....	July (−1)	.....

Most of these divergences are such as may be produced by upward trend adjustments of the data (see text below). Another type of disagreement between the chronologies, due presumably to the same cause, is recorded in Table 2 for personal income. Cloos distinguishes a specific contraction in this series in 1960–61, while the Bureau, on the basis of later data, does not. This, too, seems attributable to the fact that recent revisions of these figures have accentuated their upward trend.

NBER dates: there are twelve bracketed leads or lags in column (2) of the table and six in column (1), the difference between the two arrays being that the former refers to the unrounded and the latter to rounded data.

Rounding is by no means necessarily advisable. It will not shift the timing of those peaks that are sufficiently high locally and of those troughs that are low; but it will produce flat tops and bottoms and a bias toward late dating wherever the first differences of the series about the turning points are so small as to fall within the rounding range. This is particularly relevant for the comprehensive series, which often show retardations and gently rounded turns rather than sharply angular peaks or troughs (such as are more frequent in individual series of narrower coverage). Cloos finds that rounding mitigates the difficulty of "double bottoms"; this it may occasionally do but at the cost of creating more flat zones of indeterminacy. As for the significance of the fourth and fifth digits, little is usually known to the user and generalizations had better be avoided, but it is not clear that this information (which the compiling agencies judge worth reporting) ought to be discarded.

Among the remaining discrepancies, those that relate to the 1949 trough deserve particular consideration. The influence of a major strike occurring at the end of a business-cycle phase can produce a serious problem for the reference dating, and the 1949 revival is the classical case in point. Cloos notes that while the Bureau picks October as the reference trough, "all four major series [in his sample of aggregates] hit a low in July if one ignores the nationwide steel strike which lasted from October 1 to November 11, 1949, which was spliced onto a long coal strike." He concludes that "July is a rather good reference month, and October is much too late."[21]

But can the strike be "ignored" in such a case? If a strike occurs in the middle of an expansion or contraction, it can indeed be passed over without difficulty. This is not so, however, when the timing of the strike approximately coincides with the general business upturn, as happened in 1949. A month witnessing a major strike may nevertheless be the best choice for a cyclical trough date according to the available evidence. If thus judged to mark the beginning of a general business recovery, this month will then be selected as reference trough, although usually with more than the usual diffidence; for it is certainly very difficult to determine where the trough would have been had there been no strike.[22]

The real issue in a situation such as prevailed late in 1949 is this: Has the strike interrupted a recovery that really began earlier in the year, in July–September? Considerable improvement in output and income did, in fact, occur in these months. Some or all of this improvement, however, may have been due to stockpiling by users and producers in anticipation of the strike.[23] Recent years offer several examples of concentrated advance buying prompted by expectations of work stoppages. The problem raised by these considerations is a difficult one. It would be helpful if important comprehensive series could be estimated with sufficient confidence "strike-free." I have no such series and they would be quite difficult to construct. Cloos pro-

[21] Cloos, *op. cit.*, p. 28.

[22] On the NBER treatment of strikes in dating cycles in the affected series see Burns and Mitchell, *op. cit.*, pp. 59 and 62.

[23] Cf. Moore, "Discussion," *Proceedings of the Business and Economics Section of the American Statistical Association* (Washington, D.C., 1961), p. 34.

vides no evidence of this type to outweigh the fact, which cannot be lightly ignored, that most of the basic monthly indicators reached their lows late, in or about October.

Incidentally, it is true that the Bureau has discretion, which may sometimes be "wide," in picking the peaks and troughs in individual series; but so has Cloos or any other investigator.[24] How the specific-turn dates are identified by the Bureau's research staff has been described in detail in NBER publications, in particular in chapter iv of *Measuring Business Cycles*. For the aggregative series here discussed (as for many other "important series"), charts and tables that show the location of the NBER specific cycle turns have been published. Cloos, on the other hand, provides no aids for checking on *his* selection of turning points.[25]

Cloos's monthly reference dates are selected by taking approximately the medians of the corresponding peaks and troughs in four aggregative series (those included in Table 2 plus business sales).[26] In interpreting the results, the timing of the quarterly GNP and of "construction activity" is also quoted as a "supplementary aid." Table 3 compares Cloos's medians with the NBER monthly reference dates. The changes suggested by Cloos are as follows: shifts of ± 1 month in two cases, shifts of -3 months in two others, no change in three cases, and undecided (no change or a shift of $+2$ months) in one—for a total of eight turning points reviewed (col. [6]).

These results are virtually identical with those obtained in an earlier study by Lorman C. Trueblood.[27] The following quotation from Geoffrey Moore's comments on that study is therefore applicable:

> First the suggested revisions are fairly small. Those that the National Bureau itself made in the past, for the period 1919–38, are of a similar character. . . . By emphasizing the modest size of the revisions I do not mean to deprecate their importance, but merely to point out that while there is a zone of uncertainty about each turn it is ordinarily not large. Some of this uncertainty

[24] Erratic movements are often larger in the more specific series than in the broad aggregates, as Cloos (p. 28) observes. But, while small absolutely, they may be large enough relative to the cyclical change in the vicinity of the turn to make dating difficult. This is particularly so where cyclical turns are flat or gently rounded rather than angular. The behavior of aggregate production, employment, and income at the 1954 trough provides good illustrations (see below). There is no good reason to exclude important series from the reference-dating analysis solely because they seem more "erratic." Rather one might hope to mitigate the difficulty by examining the behavior of a larger number of series.

[25] I have compared Cloos's dates with the specific-cycle turns marked on two sets of charts for the three series in question: (1) the earlier data used by NBER for their indicator volume; (2) the latest revisions of the data available (February, 1963). I did not have the charts used by Cloos for either the unrounded or rounded, seasonally adjusted data. As noted earlier, rounding seems to be a major source of the identified discrepancies and revisions may account for some others. But my own reading of the charts I had, at least, still leaves me with some serious doubts about a few of his choices—and with no doubt at all that they would not be appropriate if he were trying to use the Bureau's procedures for dating turns.

[26] In two instances (the troughs in 1954 and 1958), the median falls between two adjoining months, and Cloos picks the later one in each pair. On one occasion (the 1960 peak), Cloos has two alternative dates that do not represent medians of the corresponding sets of individual turns (cf. Table 3).

[27] See his "The Dating of Postwar Business Cycles," *Proceedings of the Business and Economics Section of the American Statistical Association* (Washington, D.C., 1961), pp. 16–26. The only difference is that Trueblood accepts the Bureau's May, 1960, peak date with the comment that it "seems to be a reasonable choice . . . although it does represent a compromise in a rather ambiguous period" (p. 19), while Cloos suggests either May or July saying that the "exact month of the 1960 peak is very doubtful. If we are to adopt the Burns and Mitchell criterion of late dating in doubtful cases, May is too early" (*op. cit.*, p. 29).

Some further observations on the relation between the two papers are deferred to later pages.

existed when the original choices were made. This was the case, I recall, with respect to the choice between October or November, 1948, between October, 1949 or some earlier month, and between July and August, 1957.[28]

When the current (February, 1963) NBER dates for specific-cycle turns in industrial production, non-agricultural employment, and personal income are substituted for those chosen by Cloos, different medians are obtained that would match the Bureau's reference dates precisely in four instances (see Table 3, cols. [3] and [6]). The divergence at the 1949 trough is now eliminated; on the other hand, the divergences relating to the 1948 and 1954 turns would be somewhat increased.

Using current data and Cloos's method of rounding, a set of medians is derived which differs from that presented by Cloos in two instances, the 1949 trough and the 1960 peak (cf. cols. [4] and [2]). Rounding shifts three dates (in 1948, 1954, and 1957) and leaves unchanged five (cf. cols. [7] and [8]).

The discrepancies between the different chronologies that are due to data revisions bring up an important point. There is need for a review of the recent (postwar) business-cycle reference dates in the light of revised and new data now available. This has been clearly acknowl-

TABLE 3

ALTERNATIVE SETS OF MONTHLY REFERENCE DATES COMPARED WITH NBER CHRONOLOGY, 1948–61

Year of Peak (P) or Trough (T)	Cloos's Specific Turns[a]	Cloos's Median	Median Using NBER Specific Turns		NBER Monthly Reference Date	Lead (−) or Lag (+), in Months, Relative to NBER Reference Dates of		
			Unrounded Data[b]	Rounded Data[c]		Cloos's Median[d]	Median Using NBER Specific Turns	
							Unrounded Data[e]	Rounded Data[f]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1948 P.....	8, 10, 10, 11; IV, 8	October	August	October	November	−1	−3	−1
1949 T.....	7, 7, 7, 7; II, 4	July	October	October	October	−3	0	0
1953 P.....	7, 7, 7, 10; II, 4	July	July	July	July	0	0	0
1954 T.....	1, 4, 5, 8; II, 2	May	April	May	August	−3	−4	−3
1957 P.....	7, 8, 8, 8; III, 10	August	May	August	July	+1	−2	+1
1958 T.....	2, 3, 4, 5; I, 5	April	April	April	April	0	0	0
1960 P.....	1 or 7, 4, 4, 11; II, 8	May or July	April	April	May	0 or +2	−1	−1
1961 T.....	1, 2, 2, 2; I, 2	February	February	February	February	0	0	0

[a] Months are represented by Arabic, quarters by Roman numerals. The first four numbers refer to the specific turns of Cloos's four aggregative indicators; these dates are listed consecutively in chronological order. Also identified is the quarter of the turn in GNP and the month of the turn in construction activity, as noted in Cloos's comments.

[b] Based on the dates given in cols. (2), (5), and (8) of Table 2 and of the dates of turns in total business sales as given by Cloos.

[c] Based on the dates given in cols. (3), (6), and (9) of Table 2 and of the dates of turns in total business sales as given by Cloos.

[d] Intervals between the corresponding dates in cols. (2) and (5).

[e] Intervals between the corresponding dates in cols. (3) and (5).

[f] Intervals between the corresponding dates in cols. (4) and (5).

Source: See Table 2.

[28] "Discussion," *op. cit.*, p. 34. The NBER revisions for the interwar period shifted five out of eleven turns by one or two months, yielding an average shift (or "error") of ±0.6 months. The average shift implied in Trueblood's or Cloos's dates is ±1.0 months (if signs were regarded, −0.8).

edged by the National Bureau.[29] Data revisions can and sometimes do affect the reference dates significantly. This is not a new or surprising fact, and the material presented at several points in this paper confirms it strongly. Hence, even if the method of reference dating and its handling by the Bureau were perfect, errors in the dates would probably occur as a result of inadequacies of measurement and information, which are indicated by the revisions of old and the appearance of new data. It is certainly important to try to find out to what extent any errors in reference dates that may be established are due to changes in the available statistical information, because only after allowing for this type of "error" can some appraisal be made of the quality of the dating method itself.[30]

THE QUESTION OF REPRODUCIBILITY

Cloos quotes with added italics and implied doubt the following statement by Milton Friedman: "It is an empirical finding that the Bureau chronology is in fact *reasonably reproducible* and meaningful despite the failure to define precisely its meaning in terms of some single . . . measure. . . ."[31] Let us now ask whether or not the evidence of the comparative chronologies just presented is consistent with the above position.

There is admittedly no easy way to define with any precision the requirements for "reasonable reproducibility" that would appeal to one and all. But despite the inevitable vagueness of the concept, let me try to outline an acceptable test on this issue. As a minimum condition, a fair test would have to use two independently prepared chronologies for the same period; the two samples underlying these sets of dates should be selected by different people but should be comparable in size, scope, and "vintage" (i.e., they should not differ systematically because of inclusion of diverse revisions of the same series); and some common set of general rules should be followed in picking the turning-point dates in the sample series. Specifically, to test how "reproducible" the NBER chronology is, the broad working rules developed in the Bureau's cyclical analysis of individual series should be applied, or at the very least should not be grossly violated. Such rules still allow considerable leeway for judgment. But their observance would eliminate some possible inconsistencies.

A test that does not meet these requirements fails to provide a reasonable approximation to the basic conditions of the NBER dating procedure. What is then being "tested" may be just as much the differences between various revisions of certain series or between divergent conceptions of "specific cycles" as the thing we actually wish to ascertain, which is whether or not the NBER chro-

[29] Moore, *op. cit.*, p. 34.

[30] Analogously, in evaluations of the predictive quality of econometric models, attempts are made to remove the disturbing effect of imperfect statistical knowledge of the values of "predetermined" variables. Thus Henri Theil measures forecast values in terms of predicted percentage changes from the actual level of preceding year *as it was known when the forecast was made*, and compares them with observed values, i.e., percentage changes computed in the light of later data (see his *Economic Forecasts and Policy* [Amsterdam: North-Holland Publishing Co., 1958], p. 58). This is designed to "net out" from the test the influence of errors made in predicting the exogenous and lagged endogenous variables. These errors, it has been suggested, may be more serious than the errors in predicting the endogenous variables which are due to the inadequacies of the model (cf. John W. Lehman and James W. Knowles, "Comment," *The Quality and Economic Significance of Anticipations Data* [Princeton, N.J.: Princeton University Press, 1960], p. 47).

[31] Milton Friedman, "The Lag in Effect of Monetary Policy," *Journal of Political Economy*, LXIX (October, 1961), 453.

nology, given its basic rationale and premises, is reasonably reproducible.

On the other hand, if the requirements are met it is possible to evaluate the extent to which the Bureau procedure will yield different results *because of differences in individuals' judgments*. This, I think, is a meaningful and important type of test. Judgment enters into the choice of the NBER reference dates at many points.[32] Which aggregates are to be considered in identifying a reference date? How is one to evaluate their evidence versus that of the diffusion indexes? How much weight is to be attached to value series versus physical volume indexes? How much to the aggregates that incorporate a number of interpolated series versus those that do not? How adequate are the seasonal adjustments of the various series? How is one to allow for typical differences in the timing of different aggregates? For uncertainties about some of their peaks or troughs and particularities of behavior in the vicinity of the turns? For the effects of strikes, unusual weather, or other "random events"?

Business-cycle chronologies derived by independent investigators differ a great deal in time units, series, and methods used, as would be expected; hence they can hardly satisfy our criteria for comparability in any strict sense. However, they do usually have in common some of the data, since they all draw upon a limited supply of important aggregative indicators. This, plus an implicit consensus of a general kind about what constitutes cyclical movements, appears to be enough to create a large area of agreement among the different chronologies. In their 1946 book, Burns and Mitchell compared the NBER list of business cycles in the United States between 1854 and 1938 with the lists of others who have independently studied all or a substantial part of this period. These comparisons, which cover eight chronologies in addition to that of the Bureau, show remarkably few substantive discrepancies.[33] The Trueblood-Cloos chronology for the post-World War II period again resembles that of the NBER quite closely. Their list of cycles is exactly the same as the Bureau's. Not only that, but the dating of their peaks and troughs is not really very different. Indeed, the reader may be impressed more by the similarity than by the differences between these dates. One would not, after all, seriously expect them to be identical. But it may be well to note that those comparisons fail to meet our requirements for "reproducibility," because of divergences in coverage and vintage of the samples used, as well as other failures to observe the Bureau's stated procedures. Whether or not the NBER chronology is "reasonably reproducible," it is clear that Cloos does not provide any acceptable test of this issue.

The only way to get results that are alike in this area is, of course, to use identical data and procedures and to arrive at the same judgments. Cloos does obtain

[32] I am indebted to Geoffrey Moore for correspondence that brought out the variety of problems indicated in the remainder of this paragraph.

[33] See *Measuring Business Cycles*, chap. iv, esp. Sec. VI, "Dependability of the Reference Dates," Table 27 and Chart 10 with the accompanying text. Burns and Mitchell observe: "As long as statistical data remain in something like their present state, theoretically distinct methods of dating business cycles—each used in a thoughtful and discriminating fashion—are reasonably certain to merge in practice" (p. 92). Even with the present, much richer, supply of data, this statement seems to me to be essentially valid. It is also worth noting that there is much less uncertainty today about the typical length of business cycles than there was before the dissemination of findings based on statistically derived chronologies, primarily that of the Bureau. (It was then claimed by some that the average duration of the cycle is four years, by others that it is seven or ten years.)

the same dates as those Trueblood derived in an earlier study, apparently because of a close agreement of this sort.

ON STRAW MEN, SUSPECTS, AND GUILT BY ASSERTION

In my opinion, the measures presented by Cloos also fail to *prove* another point he is trying to make, namely, that "it is not possible to pinpoint a single month as the peak or trough of a cycle in many cases."[34] The logic of some of his inferences seems to be that unless all his four series show coincident turns, the dating is impossible. Thus he concludes: "The 1957 peak is impossible to spot in a single month, but the Bureau's peak is certainly too early." This is because "Sales hit a high in July, but the other major series did not do so until August."[35]

However, if Cloos's contention is read to mean that single-month dating is not always equally dependable and that sometimes its dependability is much less than one would desire, then of course he is right. But whose position is he contesting here? Surely not that of the Bureau. The NBER publications from which he quotes deal at length with the difficulties and uncertainties encountered in developing a chronology of business-cycle turns in monthly terms, and, obviously, the Bureau's own revisions of its reference dates imply recognition of the possibilities of error.[36]

Cloos similarly sets up and knocks down a straw man when he declares that "in cases when this [i.e., the pinpointing of a single month as the peak or trough of a cycle] can be done the Bureau is sometimes wide of the mark by as much as three months."[37] He himself provides quotations that show that the Bureau was and is well aware of this possibility and repeatedly warns the reader about it.[38] But the important point in this context is simply that a usable reference chronology is worth the effort. Imperfect as it is, we are better off with it than without it (a phrase that might again remind one of money, though it also applies to many much less needed but useful things). To be sure, the proper objective is to get the "best possible," and I see no evidence that the Bureau's efforts are not directed toward that goal or that they keep missing it.

Another of Cloos's announced purposes is to show "that the Gross National Product and the FRB industrial production index are usable measures of general business activity and that peaks and troughs in these series are to be preferred to the Bureau's peaks and troughs."[39] Here again what he sets out to prove either requires no proof or gets none. After what was said earlier on the timing of GNP and industrial production, little needs to be added here. GNP and the FRB index certainly belong to the most "usable" measures of business activity; few will dispute this and the Bureau's staff, to my knowledge, definitely has not. But the superiority of the turning points in GNP or the FRB index to the NBER chronology has yet to be demonstrated.

That the Bureau's reference dates are subject to misuse is no doubt true. So are most methods and even the firmest findings in economics, and so would be any

[34] Cloos, *op. cit.*, p. 14.

[35] *Ibid.*, p. 29.

[36] See, e.g., *Measuring Business Cycles*, chap. iv, Sec. V; Moore in chap. v of *Business Cycle Indicators*, p. 146.

[37] Cloos, *op. cit.*, p. 14.

[38] *Ibid.*, p. 17 (these quotations are from *Measuring Business Cycles* but more recent NBER statements made in the same spirit are not difficult to find).

[39] *Ibid.*, p. 14.

alternative set of dates, even if it came much closer to perfection than has ever been claimed on behalf of the NBER set. But it must be said that Cloos has little success in providing illustrations of the misuses. He apparently sees no harm in the practice of shading the reference contractions on charts because "an observer can follow the movements in the data directly," but notes that "this is more difficult when charts or tables show a number of cyclical movements in a series using the reference quarters or months as a starting point."[40] But even if it were "more difficult" to trace the movement of a series in the latter case, this need not make the approach either irrelevant or misleading. Cloos refers particularly to the charts (labeled "Comparisons of Reference Cycle Patterns") which are presented regularly in the Census Bureau's monthly *Business Cycle Developments* (*BCD*) after having been introduced in recent cyclical studies by the National Bureau. In each case, these charts are shown along with the charts of the corresponding time series proper. Moreover, graphs of percentage changes in the same series *as measured from their own peaks or troughs* are also included in the same publication (see the "Comparisons of Specific Cycle Patterns," a regular feature of *BCD*). How can any careful user be misled by any of the charts Cloos objects to if he can always cross-check on its significance by turning a page or two and looking up another chart?

Confidence in the integrity and competence of those research organizations and data-compiling agencies whose reputation is generally high may and probably does induce some users to accept the statistical products of these sources uncritically. This, everyone will agree, is regrettable. For the most part, however, it is difficult to put the blame for this on the source. Many detailed warnings against misuse of the GNP-accounts data, for example, are issued periodically by the National Income Division of the Department of Commerce, and ignored just as often by various consumers of these statistics. The Federal Reserve Board, similarly, has often urged that caution be exercised in using its own data. Surely, the National Bureau has not been less explicit in pointing to the limitations of its reference chronology (as well as of many results it has obtained with the aid of this measurement device).

I conclude that each of the major theses Cloos advances at the beginning of his paper remains unsubstantiated. But one point that seems important to me is well illustrated by Cloos's article, and that is the need for more detailed and up-to-date information on the NBER procedures for determining the reference-cycle dates. About this more will be said in the concluding section of this paper.

Let me add a comment on a matter of form. A conspicuous feature of the article reviewed is a large number of quotations from Mitchell and Burns, etc., with copious underlining by Cloos. Despite a disclaimer, this method frequently results in quoting small excerpts out of context. Many of the quotations represent qualifications and warnings regarding the pitfalls of reference dating. In the original texts, they are a part of balanced presentations that make it clear that the history of business cycles also provides many examples of well-defined turns. Their use by Cloos, however, tends to suppress this fact. The method, which I would call one of "implicit criticizing," seems to produce suspects; it can never furnish any proofs of guilt.

[40] *Ibid.*

ON FLAT TOPS AND BOTTOMS

The NBER practice in those cases where the turning points in different aggregates are widely scattered and the turning zones are relatively flat is to place the reference turn toward the close of the transition period. The Bureau attempts to treat the peaks and troughs symmetrically in this respect. If movements having the appearance of "flat bottoms" or "flat tops" were often followed by renewed contractions and expansions, respectively, and if such retardations were about equally likely in either phase, then both the procedure of "late dating when in doubt" and its symmetrical application would have considerable justification. The objective of reference dating is to ascertain when the forces of expansion gained ascendance over those of contraction in a revival (and vice versa in a recession), which is done by identifying, in a sense, the "center of gravity" of a turning-point zone; but occasionally the scatter of the turns is such as to make this determination very difficult. One can then tell with more ease and confidence when, say, an expansion has definitely begun than when the contraction was overcome, and this is the basis of the Bureau's compromise procedure in such cases.

Trueblood observes that the NBER methods of cyclical analysis were developed in the pre–World War II times "when a recession could become a depression and when recoveries could be reluctant and halting." Since the 1930's and after the war, however, structural changes in the economy and policy developments were such that "recessions have been short and relatively mild and steady recovery has been assured." Trueblood, therefore, favors earlier trough dating in a "flat-bottom" case and "treating a subsequent period of either little change or very slow rise as part of the cyclical recovery."[41] Cloos expresses the same opinion in strong and unqualified words; according to him it is "altogether wrong to date a recession low toward the end of a plateau rather than the beginning." But he argues for a different treatment of the downturn, believing that there "the Bureau's rule of late dating is valid. There have been instances such as 1947, 1951, and 1956 when plateaus or slight declines in activity were the prelude to a resumption of a business expansion."[42]

A rule such as that of late dating was never meant to be applied strictly or mechanically; it may have to do when, and as long as, positive information to resolve doubt is lacking, but it is always subordinate to the economic and statistical analysis of the interrelated processes that participate in the business-cycle upturn (or downturn).[43] A rule that the trough should always be placed at the beginning of a plateau would, I think, be as "wrong" as the rule that it should always be placed at the end: either precept would be unduly restrictive. However, shunning such extremes, the arguments of Trueblood and Cloos on this point are important and deserve to be carefully considered. In considering the matter of symmetrical treatment of peaks and troughs, one point to be weighed is that symmetry does tend to produce unbiased estimates of the durations of expansions relative to contractions.

[41] Trueblood, *op. cit.*, p. 19. It must be said that developments during 1962 raise some question whether reluctant or halting recoveries have indeed been banished.

[42] Cloos, *op. cit.*, p. 20.

[43] Consider the following passage from *Measuring Business Cycles*, starting on p. 77 (italics added): "When we first made a reference scale of business cycles, we had to rely on vague knowledge concern-

ON LEADING INDICATORS

Of Cloos's digressions, it will be sufficient to mention just one. According to Cloos, "It is interesting that only the 'coincident' indicators bear the adjective 'roughly'—one supposes this is intended to give expression to the Bureau's view that there is no useful aggregate of economic activity."[44] This is indeed a puzzling imputation. In all NBER studies, "rough coincidence" is a technical term denoting a measure that is contrasted with that of "exact coincidence." Both are precisely defined. "Exact coincidence" precludes any measurable lead or lag in the given individual observation or average. "Rough coincidence" includes exact coincidences and leads and lags of one, two, and three months.[45]

Heedless of this definition, Cloos proceeds to add the "leads," "lags," and "rough coincidences" for the "leading indicators" of the Bureau. He concludes that these twelve series "actually led peaks and troughs from 1854 on in only 209 of 344 observations" and that "One might have expected better results on the basis of chance alone."[46] Unfortunately, by adding the above three categories to obtain 344 observations, he counts the short leads and lags twice. What should be added is leads, lags, and *exact* coincidences. This gives a total of *255* timing observations, out of which 209 are leads.[47] The probability calculus is not so simple here because of lack of independence, but certainly the correctly computed proportion of leads, which is 82 per cent, has rather different implications than Cloos's wrong result of 61 per cent.

In addition to criticizing the indicators for not leading the reference dates frequently enough, Cloos also criticizes them for sometimes leading "arbitrary points in time rather than *the* cycle turning point."[48] Thus a date that deviates from Cloos's choice by one to three months in either direction is implicitly defined as "arbitrary"—whatever may be the total net weight of other evidence in its favor. This is difficult to accept. Suppose one had full confidence in Cloos's analysis of postwar turning points and decided that, in each instance of a divergence, his date was exactly right and that selected by the Bureau wrong: even this would not have been sufficient to demonstrate either the "arbitrariness" of the NBER dates or the alleged failure of leading indicators. Occasional shifts or "errors" in the reference dates of the order of one to three months (see Cloos's results reproduced in Table 3, col. [6]) would be definitely insufficient to deprive the series classified by the Bureau as "leading" of their tendency to lead at cyclical revivals and recessions. The mean lead of these series in the period 1948–58, for example, was eleven months; leads of four months and over accounted for 88 per cent of all observed leads and for 79 per cent of all timing comparisons made for this group.[49] It seems safe to

ing the cyclical behavior and the economic significance of different series. . . . In many cases the turning points of different series were bunched so closely that we could not go far astray. But there were cases in which the turning points were widely scattered, and others in which they were concentrated around two separate dates. *If there was little else to guide us*, we placed the reference turn toward the close of the transition period."

[44] Cloos, *op. cit.*, p. 27.

[45] See, e.g., *Business Cycle Indicators*, p. 57.

[46] Cloos, *op. cit.*, p. 27.

[47] We are using here the new (1960) list of leading indicators as recorded in *Business Cycle Indicators*, Table 3.2, pp. 56–57. It is to these measures that Cloos refers.

[48] Cloos, *op. cit.*, p. 32.

[49] See G. H. Moore, "Leading and Confirming Indicators of General Business Changes," chap. ii, of

conclude that it would take much larger "errors" than those Cloos believes to have found in order to invalidate the NBER leading indicators—and the errors would have to be due to a systematic post-dating bias.

It should be noted that timing measures for the other two groups of indicators, the coinciders and the laggers, contribute additional evidence that is important for an appraisal of the performance of the leaders. If the high proportion of leads for the latter group were due to a selection of tardy reference dates, then this would show up in a bias toward too early timing on the part of the other indicators. That is, the coinciders and laggers would have lower proportions of rough coincidences and lags, and, in the case of sufficiently large errors, would have to be disqualified on this ground. The observed systematic differences in the distribution of leads and lags among the three groups of indicators testify against the existence of such a bias.

AN ILLUSTRATION OF REFERENCE DATING: THE 1954 TROUGH

Several major considerations involved in selecting a reference date can be illustrated by examining the case of the 1954 trough, which is one of the most difficult ones on the record. Cloos terms this trough "highly indefinite" but thinks that the NBER choice, August, is late by three months (see Table 3) or perhaps more.[50] Of the postwar turns, only the 1949 trough seems to present a problem of comparable difficulty.

The National Bureau first selected August, 1954, as a reference trough early in the winter 1954–55. The evidence then on hand suggested this original choice and gave it considerable support. The troughs in the important comprehensive series, as identified from seasonally adjusted data available in January, 1955, are listed in Table 4. I shall not try to reproduce the considerably larger body of data on which the Bureau's selection was based.[51]

Revised and new data examined at the Bureau later in 1955 disclosed a wider dispersion of the 1954 troughs but gave no sufficient reasons to refute the original choice of the reference date. Series on employment and unemployment continued to support that choice strongly, and so did the related series on man-hours that were constructed at a later time.[52] Labor income in non-farm commodity-producing industries also turned up in the third quarter of 1954 (September), although total personal income reached its trough early in the year. Diffusion indexes were constructed from (*a*) non-agricultural employment and (*b*) industrial production series, by industries, and these disclosed that August was the first month

Business Cycle Indicators I, pp. 52–53 (Table 3.1). These measures pertain to the eight leading series of Moore's 1950 list, and hence to a period not covered by the data at the time the series were selected. Cf. also the median leads for the twelve leaders in the new (1960) list, *ibid.*, Table 3.2.

[50] Cloos, *op. cit.*, p. 29.

[51] The first published designation of August, 1954, as a tentative date of the trough is found in Moore's statement for the *Thirty-fifth Annual Report of the National Bureau* (May, 1955), along with some evidence bearing on the chronology of business contractions in 1948–49 and 1953–54. For an assessment of the situation when data through August and September were the latest available, see Moore's paper before the Conference on the Economic Outlook, University of Michigan, November 11 and 12, 1956, published in the *Michigan Business Review*, January, 1955, pp. 1–8. See also *Business Cycle Indicators*, pp. 77–79.

[52] The evidence of employment series is sometimes seriously affected by strikes, but this apparently does not apply in the present case. Strike activity in 1954 was relatively light, even at its highest level in the summer. Man-hours in the non-farm sector reached their troughs in July and August both for persons with jobs and for those at work.

when more than half the industries covered were expanding. The cumulated diffusion index based on the National Bureau's eight coincident indicators also suggested that the balance did not shift toward expansion until August. Early revisions did shift backward in time the upturns in some important series, including GNP. However, an analysis of thirty-one aggregative series and nineteen diffusion indexes completed at the Bureau in November, 1955, by Alexander Pitts, still pointed to August, 1954, which was the shifted to the third quarter, where it remained from November, 1954, through June, 1955. After July, 1955, the trough shifted to the second quarter, after July, 1956 (for two years) to the first; and now (since July, 1958) it has come back again to the second.

The moral to this story is clear: the GNP figures would have hardly made a dependable guide to reference dating on this occasion. Changes in GNP were exceedingly small between the first and the third quarter of 1954. This was a period

TABLE 4

SPECIFIC CYCLE TROUGHS IN SELECTED COMPREHENSIVE SERIES, 1954, DATA AVAILABLE IN JANUARY, 1955

Series	Date of Trough
1. Personal income	April, 1954
2. FRB index of industrial production[a]	August, 1954
3. GNP in current dollars, quarterly	III, 1954
4. National income, quarterly	III, 1954
5. Non-agricultural employment, BLS[b]	September, 1954
6. Sales, manufacturing and trade	October, 1954

[a] The index available in 1955 covered manufacturing and mining; in the 1959 revision, it was broadened to include utility output of electricity and gas. About the effects of this and other changes, see *Industrial Production—1959 Revision* (Washington, D.C.: Board of Governors of the Federal Reserve System, July, 1960).

[b] A count of full-time and part-time employees on payrolls, based on reports of non-farm establishments.

Source: U.S. Department of Commerce; Board of Governors of the Federal Reserve System; U.S. Department of Labor. Based on data given in *Economic Indicators*, January, 1955, prepared for the Joint Committee on the Economic Report by the Council of Economic Advisers (Washington, D.C., 1955).

median of the turning-point distribution for this sample. The selection of August and the third quarter of 1954 as the monthly and quarterly reference troughs was therefore confirmed by the Bureau.[53]

The history of the revisions of GNP is interesting. Shifts in the date of the 1954 trough in GNP occurred no less than six times between July, 1954, and July, 1958 (Table 5). First the lowest standing of GNP was briefly recorded in the first quarter, then (from August through October, 1954) in the second; with the publication of the third-quarter figure, it of strong cross-currents and a sluggish over-all change. According to present data, GNP in current dollars, as derived from expenditure accounts, declined just 0.33 per cent between the first and the second quarter; by so little, indeed, that the mere omission of the "statistical discrepancy" is sufficient to change the result into an increase (of less than 0.2 per cent) in the "income accounts" variant of the GNP series. Subsequently GNP increased less than 1 per cent between the second and the third quarter of the year—however measured. It was not until the last quarter of 1954 that GNP turned up decidedly, by more than 2 per

[53] *Thirty-sixth Annual Report* (New York: National Bureau of Economic Research, May 1956), p. 27.

cent. The relative changes in the constant-dollar GNP were still smaller.

After recent upward revisions, the FRB production index and the diffusion index based on its components now appear to have turned in April and March rather than August; and the NBER diffusion index of the roughly coincident indicators now suggests April rather than August as the trough. This strengthens the case for earlier dating of the 1954 upturn. However, plotted on semilogarithmic scales, curves of the currently available data for such comprehensive series as industrial production or personal income (as well as GNP) are still very flat in the vicinity of their 1954 troughs. The upward tilts imparted to them by the recent revisions are slight. As Lorman Trueblood rightly puts it, "The question whether the period from the spring to late summer 1954 should be treated as a protracted 'bottoming out' period and part of the recession, or as the beginning of cyclical recovery, thus balances on a fairly narrow statistical margin."[54]

As noted earlier, cyclical reversals that are widely diffused and of the "gently rounded or flat" type are particularly difficult to date; far more so than the "angular" variety of relatively concentrated turning-point zones. Imperfections of data and erratic events affect the former much more than the latter. The 1954 episode illustrates well the "flat-bottom" problem. In selecting August the Bureau picked a late low month, which had the advantage of pointing to a definite upturn; the uncertainty attaching to an earlier choice would have been considerably greater.

The question now is whether the subsequent revisions of the data have made a shift of this reference trough to an earlier month advisable. Trueblood answers it affirmatively, saying that "whereas the preliminary figures pointed to an August

TABLE 5

GROSS NATIONAL PRODUCT, QUARTERLY, IV, 1953—I, 1955, REVISIONS PUBLISHED BETWEEN JULY, 1954, AND JULY, 1958

DATE OF PUBLICATION AND SOURCE[a]	CURRENT VALUE OF GNP, QUARTERLY AT ANNUAL RATE, SEASONALLY ADJUSTED[b]					
	IV, 1953	I, 1954	II, 1954	III, 1954	IV, 1954	I, 1955
July, 1954 (EI)	363.5	357.2	356.0[c](T)			
August and September, 1954 (EI)	360.5	355.8(T)	356.0			
October, 1954 (EI)	360.5	355.8(T)	356.0	356.0[c]		
November, 1954, through January, 1955 (EI)	360.5	355.8	356.0	355.5(T)		
February, 1955, through April, 1955 (EI)	360.5	355.8	356.0	355.5(T)	362.0	
May and June, 1955 (EI)	360.5	355.8	356.0	355.5(T)	362.0	370.0
July, 1955 (EI, SCB)	364.5	358.3	357.6(T)	358.8	367.1	375.3
July, 1956 (SCB)	357.5	357.6(T)	358.5	359.4	367.1	377.3
July, 1957 (SCB)	363.2	358.1(T)	358.7	360.0	367.7	379.0
July, 1958 (SCB)	361.0	360.0	358.9(T)	362.0	370.8	384.3

[a] EI: *Economic Indicators* (Council of Economic Advisers); SCB: *Survey of Current Business* (Office of Business Economics).
[b] The lowest standing of GNP in each line is denoted by T in brackets (trough).
[c] Preliminary estimate by Council of Economic Advisers.

[54] Trueblood, *op. cit.*, p. 18.

trough, the revised figures point to May."[55] There is no need for me to make a specific choice here, and I would not try to make one without some further analysis. But it is clear that the effects of the revisions have in fact the indicated direction.

[55] *Ibid.*

A new comprehensive description of the NBER method of reference dating is now being prepared at the Bureau, with particular attention to the peaks and troughs of the postwar business cycles. This work will involve a systematic review of the recent chronology and will produce such revisions of it as may be required.