

NBER WORKING PAPER SERIES

REMEASURING BUSINESS CYCLES

Christina D. Romer

Working Paper No. 4150

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

August 1992

I am grateful to Geoffrey Moore, David Romer, Glenn Rudebusch, Richard Sutch, and Victor Zarnowitz for helpful comments and suggestions and to the National Science Foundation and the Alfred P. Sloan Foundation for financial support. This paper is part of NBER's research programs in Economic Fluctuations and Development of the American Economy. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #4150
August 1992

REMEASURING BUSINESS CYCLES

ABSTRACT

This paper evaluates the consistency of the NBER business cycle reference dates over time. Analysis of the NBER methods suggests that the early turning points are derived from detrended data, while the dates after 1927 are derived from data in levels. To evaluate the importance of this and other changes in technique, the paper derives a simple algorithm that matches the postwar NBER peaks and troughs closely. When this algorithm is applied to data for 1884-1940, the new dates systematically place peaks later and troughs earlier than do the NBER dates. Using the new business cycle chronology, recessions have not become shorter, less severe, or less persistent between the pre-World War I and the post-World War II eras. Expansions, however, have become longer.

Christina D. Romer
Department of Economics
University of California, Berkeley
Berkeley, CA 94720
and NBER

I. INTRODUCTION

The business cycle reference dates for the United States produced by the National Bureau of Economic Research (NBER) occupy a unique position in the fields of macroeconomics and economic history. These dates, which are intended to show the peaks and troughs of economic activity from the mid-1800s to today, play a crucial role in forming impressions about macroeconomic fluctuations. For example, when a new peak or trough is announced today it is considered front page news and is often thought to have substantial political ramifications. Similarly, the NBER reference dates for the late 19th and early 20th centuries are often taken as the definitive summary of whether conditions were good or bad in some particular era.

In addition to affecting impressions of certain years both in the past and the present, the NBER reference dates are a much used tool of macroeconomic analysis. They are frequently employed, for example, to analyze how some particular variable behaves in relation to the business cycle. A variable that has no obvious relation with the NBER turning points is likely to be dismissed as a causal factor, while one that is highly correlated may be considered a candidate for explaining macroeconomic fluctuations.

Finally, the NBER reference dates for the United States have been influential in forming impressions about possible changes in business cycles over time. For example, the NBER dates are often used to measure the length of expansions and contractions in certain eras. Since the NBER dates for the United States show that expansions have become longer over time and contractions have become shorter, many economists have inferred that economic fluctuations have become less severe between the prewar and postwar eras.¹

A. Issues

Part of the reason that the NBER reference dates have been so influential is that they are simply very convenient. They provide a quick shorthand that economists can use to summarize a very complex

phenomenon. More fundamentally, the NBER dates have been influential because they are thought to be reliable. The amount of work that went into their development is extremely impressive. Burns and Mitchell's seminal study Measuring Business Cycles (1946), in which the NBER methodology is described and developed, is surely one of the most cited and respected books in American macroeconomics.

While Burns and Mitchell's contribution to business cycle dating is indeed monumental, it is nevertheless reasonable to investigate the accuracy of the NBER reference dates and particularly the consistency of these dates over time. Statistical techniques and understanding of the nature of economic fluctuations have advanced greatly in the 45 years since Measuring Business Cycles. It is possible that these advances have caused the procedures used to select the early reference dates to have been improved upon. In addition, even in the absence of such advances, it is possible that unintended inconsistencies have crept into the NBER dating procedures over the long period that the NBER has been setting reference dates.

Perhaps the strongest urging for such questioning of the NBER reference dates comes from Burns and Mitchell themselves. In Measuring Business Cycles they state:

This is not to say that the reference dates must remain in their present state of rough approximation. Most of them were originally fixed in something of a hurry; revisions have been confined mainly to large and conspicuous errors, and no revision has been made for several years. Surely, the time is ripe for a thorough review that would take account of extensive new statistical materials, and of the knowledge gained about business cycles and the mechanics of setting reference dates since the present chronology was worked out (1946, p. 95).

Burns and Mitchell had planned to carry out this thorough review, "but this project had barely started when the investigators placed in charge were drawn into war work. For the time being, therefore, we must put up with a reference scale that requires extensive reworking" (1946, p. 95). Such reworking, however, has never occurred.

B. Overview

In this paper I examine the consistency of the NBER reference dates over time. Section II compares the methods that the NBER uses today to date cycles with the methods employed by early NBER researchers. I find that the NBER reference dates for the cycles before 1927 were chosen long before the modern procedures described in Measuring Business Cycles were established. As a result, the methods used to date the early cycles are quite different from those used in the postwar era. Most importantly, whereas modern reference dates are based on data that includes the secular trend, the pre-Depression reference dates appear to be derived primarily from detrended data. This change is likely to make contractions appear longer in the prewar era than in the postwar era because any series with an upward trend will tend to peak earlier and trough later when examined in deviations from trend instead of levels.

Sections III and IV present statistical evidence on the importance of these changes in technique over time. In Section III I derive a simple algorithm for identifying cycles and choosing turning points from monthly data on industrial production that yields dates that match the postwar NBER reference chronology very closely. In Section IV I use this same algorithm to choose turning points in the prewar index of industrial production recently compiled by Miron and Romer (1990). The prewar turning points deduced using the algorithm are systematically different from the NBER reference dates: many of the new prewar peaks are several months later than the NBER peaks and many of the new troughs are several months earlier.

Section V analyzes what the new, more consistent, dates reveal about possible changes in the business cycle over time. Most obviously, the new dates affect the comparison of the average duration of recessions in the prewar and postwar eras. Whereas the NBER reference dates show a dramatic decline in the length of contractions over time, the new dates that I derive show an increase in average duration of nearly three months

between the pre-World War I and the post-World War II eras. The new dates, however, confirm the traditional finding that cycles are more frequent in the prewar era. Analysis of the new dates, in conjunction with the industrial production series from which they are derived, shows that the average output loss associated with recessions is quite similar before World War I and after World War II, and that the time it takes output to return to the previous peak level is slightly shorter in the prewar era than today.

II. CHANGES IN NBER DATING METHODS OVER TIME

A. Modern NBER Dating Procedures

According to Moore and Zarnowitz (1986), the procedures that the NBER currently uses to date business cycles were laid down in more or less their final form in Measuring Business Cycles. And, indeed, Moore's (1983) and Zarnowitz and Moore's (1983) descriptions of the actual working of the NBER Committee on Business Cycle Dating are consistent with the principles spelled out in this work. In Measuring Business Cycles, Burns and Mitchell draw a distinction between "specific cycles," which are the peaks and troughs in a particular series, such as industrial production or bank clearings, and "reference cycles," which are the peaks and troughs for the economy as a whole. The reference cycle dates are what we typically think of as the turning points of business cycles. These reference dates do not necessarily reflect the extremes of any one aggregate series, but rather reflect a consensus of the turning points in many specific series.

Specific Cycles. The procedures that Burns and Mitchell suggest for identifying specific cycles and their turning points involve many steps and guidelines. The first is that it is best to work with data that has been seasonally adjusted, but not detrended. To identify cycles in seasonally adjusted, but otherwise "raw" series, Burns and Mitchell look for "well-defined movements of rise and fall" (1946, p. 57). That is, they seek to

identify actual declines in the series, not mere slowdowns in growth.²

To decide which rises and falls are significant enough to be classed as specific cycles, Burns and Mitchell use a combination of a duration rule and a minimum amplitude rule. For a fluctuation to count as a specific cycle the duration must be "at least 15 months, whether measured from peak to peak or from trough to trough" (1946, p. 58). It must also be less than ten or twelve years in length.³ The amplitude rule takes the somewhat circular form that "the lower limit of the range of amplitudes of all fluctuations that we class confidently as specific cycles is our rough guide in deciding whether any doubtful fluctuation . . . is well enough defined to be accepted as a specific cycle" (1946, p. 58).

For some cases the identification of specific turning points, once a given movement is classified as a cycle, is straightforward. If the highs and the lows of the series are unique and obvious, the months in which those extremes occur are taken as the turning points. But in other cases, the identification of turning points is more complicated. For example, if the series flattens out around the peak or trough, Burns and Mitchell use the rule that the "latest month in the horizontal zone is chosen as the turning date" (1946, p. 58). If there are multiple peaks or troughs, Burns and Mitchell tend to date the turning point at the latest extreme, provided that there has not been a significant decline before the latest peak or a significant rise before the latest trough (1946, p. 58).

Reference Cycles. Burns and Mitchell's description of how reference cycles are dated is much less precise than their discussion of specific cycles. They emphasize that reference cycles should not be identified by choosing the turning points in one aggregate series because, while a cycle must show up in a measure of aggregate economic activity, there could be movements in some aggregate that are driven by only a few series. For this reason, Burns and Mitchell prefer to look at the turning points in many specific series and check for coherence.

How such coherence is determined and particular dates for reference

cycles actually chosen appears to be deliberately vague. Burns and Mitchell emphasize that they do not just average the turning points in various specific series. Nor do they have some particular weights attached to various series. Rather, they seem to rely on subjective judgment and an internal weighting scheme for deciding which series to use as their main guide. They refer to refining the "approximate dates by arraying the cyclical turns in the more important monthly or quarterly series" (1946, p. 77). Moore and Zarnowitz elaborate that, at least in the modern application of Burns and Mitchell's procedures, the series that deserve the most attention are ones that are "more comprehensive, more significant economically, more adequate statistically" (1986, p. 747).

B. Early NBER Dating Procedures

Sources. While Measuring Business Cycles contains much information about how reference cycles should be dated in the future, it gives only a very brief description of the procedures that were actually used to derive the NBER reference dates for the late 1800s and early 1900s. The reason for the vagueness of Measuring Business Cycles about early procedures is probably the fact that the reference dates for the cycles before 1927 were set long before Measuring Business Cycles was written. Burns and Mitchell, in the same section in which they extol the need for revamping the reference dates, explain in a footnote that the "American reference dates through 1927 have been allowed to stand as published in 1929" (1946, p. 95)⁴

The earliest source for the NBER reference dates appears to be an article in the NBER News-Bulletin for March 1, 1929 entitled "Testing Business Cycles." This article, which is unsigned, is a summary of a (supposedly) forthcoming work by Mitchell. The article itself is listed in a bibliography of Mitchell's writings compiled by Burns (1952, p. 357), suggesting that Mitchell is the author.⁵ Two other roughly contemporaneous sources for the NBER reference dates are Mitchell's entry on business cycles in the Encyclopedia of Social Sciences published in 1930

and Mitchell's chapter entitled "A Review," in Recent Economic Changes published in 1929.

The book that Mitchell was working on in 1929 was never published in its original form. In late 1933 Mitchell took on Arthur Burns as a collaborator, and the work was revised and eventually published as Measuring Business Cycles in 1946.⁶ The fact that the reference dates for the cycles before 1927 were set by Mitchell before 1929 implies that Burns was not involved in the original dating of the early reference cycles. This is important because Mitchell, in his preface to Measuring Business Cycles, states that Burns "instituted a searching critique of our methods and rigorous tests of our findings, out of which came many improvements in our conceptions and procedures" (1946, p. vii). Thus, it would not be surprising if the methods used to set reference dates were quite different before and after Burns's collaboration.

Given that it was Mitchell who ultimately set the NBER reference dates for the years before 1927, the obvious source to consult for the NBER procedures is Mitchell's book Business Cycles: The Problem and Its Setting (1927), which was published less than two years before the first discussion of the NBER reference dates. In this book Mitchell does not actually set reference dates. However, there are extensive discussions of the methods one would use to isolate cycles and of the data series that the NBER was currently analyzing.

Methods. All of the early works on the NBER reference dates refer to two major sources used to identify peaks and troughs: Business Annals and business indexes. "Testing Business Cycles," for example, says "a set of 'reference dates', based on the National Bureau's Business Annals supplemented by business indexes, is made for each country covered to show the month and year when economic revivals and recessions occurred" (1929a, p. 2). Because the description of the derivation of the early reference dates in all the sources does not go beyond such brief statements, it is necessary to back out the procedures and criteria that the researchers at the NBER

used for dating early business cycles.

The Business Annals to which all these works refer is a compendium of contemporaneous opinion on the state of the economy in various countries that was collected by Willard Thorp and published by the NBER (1926). This work provides a summary of what periodicals such as The Commercial and Financial Chronicle or Dun's Review said about when periods of recession and revival occurred in the 1800s and early 1900s.⁷ Chart VI of Mitchell's long introduction to Business Annals (1926, pp. 94-95) shows graphically the periods of prosperity, recession, depression, and revival that Thorp identified from the business press.

The Annals appear to have been used mainly for deciding roughly when cycles were rather than for establishing the actual reference dates. Burns and Mitchell report that the Annals were used to "write down an interval within which a cyclical turn in general business probably occurred" (1946, p. 77). The importance of the Annals for deciding whether or not a cycle occurred can be seen in the fact that there are no cycles identified in the Annals that are not in the reference scale, and no cycles in the reference scale that are not identified in the Annals.

There is much less correspondence between the actual dates of quarterly turning points deduced from the Annals and the NBER reference dates.⁸ Of the 24 peaks and troughs dated by the two sources for the period 1887 to 1924, the two agree perfectly in only ten instances. The average difference between the two sets of dates for the remaining 14 turning points is two quarters, with the largest difference being four quarters. This lack of correspondence is equally noticeable in periods for which the descriptive summaries of the Annals give a very exact opinion about a turning point. For example, the Annals for 1891 read "dullness continues until August, when revival sets in" (1926, p. 136), but the NBER reference date for the trough of this cycle is May 1891.

The fact that the NBER reference dates differ quite markedly from the dates given in Business Annals suggests the business indexes to which

Mitchell often refers must have played a significant role in the derivation of the actual NBER peaks and troughs. Business Cycles (1927) provides crucial information on which of the available statistical series are likely to have been used to set reference dates. The two series that receive the most attention in Business Cycles are the A. T. and T. Business Index and the Snyder Clearings Index of Business. For example, many graphs show only these two series. Furthermore, at various points in the text, Mitchell refers to these two series as "the most significant of the statistical indexes which run back to 1875" (1927, p. 422) or "the leading American . . . indexes of general trade which cover considerable periods" (1927, p. 367). Thus, it seems likely that these two series are the main ones that Mitchell consulted in choosing reference dates for the late 1800s and early 1900s.

Both of these series were intended to be comprehensive indicators of general business. The A. T. and T. index includes such diverse series as pig iron production, wholesale prices, and bank clearings. The Snyder clearings index is based on bank clearings outside New York City deflated by a weighted average of various price indexes. A crucial feature of both series is that they were detrended by their creators.⁹ That Mitchell worked with detrended series in setting the early reference dates is consistent with his statement in 1930 that the most promising way to identify business cycles is to begin by "eliminating the secular trend of a series" (1930, p. 94).

In Business Cycles, Mitchell lists the "Dates of the Troughs and the Crests of ... Five Monthly Indexes of Business Activity" (1927, Table 14, p. 335). A comparison of these troughs and crests with the underlying data shows that Mitchell is simply reporting the actual extremes in the various series. When there were multiple crests and troughs of exactly the same magnitude, Mitchell lists both. Table 1 reproduces the dates of crests and troughs in the A. T. and T. index and the Snyder clearing index, along with the NBER reference dates.

Based on these troughs and crests in the A. T. and T. and Snyder indexes, it is possible to deduce rules for choosing turning points that match

the NBER reference dates quite closely. For troughs, the rule is to choose the later of the absolute troughs in the A. T. and T. and Snyder series as the reference date. This rule fits seven of the 12 pre-1927 reference troughs exactly. Four other NBER reference troughs are one month after the later of the two series. For peaks, the rule is to choose the absolute peak in either the A. T. and T. or Snyder series as the reference date, without any systematic tendency to choose the earlier or later of the two series. Eleven of the twelve NBER peaks correspond to the absolute peak of one of these two series.¹⁰ The NBER reference peaks generally correspond to the peaks of the Snyder series through 1900 and the peaks of the A. T. and T. series thereafter. No simple rule, however, matches the NBER reference dates perfectly. This is not surprising because even if early NBER researchers followed approximately these criteria, there is no reason to expect them to have applied the rules in a mechanical fashion.

That the early NBER reference dates appear to have been set largely on the basis of just a few detrended business indexes may seem to contradict the numerous references in Measuring Business Cycles to the hundreds of series analyzed by the NBER. The resolution of this seeming contradiction is that the reference dates were an input to the analysis of many other series, not the final result of this analysis. As Burns and Mitchell note:

To learn how different economic processes behave in respect of business cycles, their movements must be observed during the revivals, expansions, recessions, and contractions in general business activity. Before we can begin observing we must mark off these periods (1946, p. 24).

Thus, it is not surprising that the early reference dates are derived from only a limited amount of information.¹¹

C. Possible Implications of the Changes in Technique

A comparison of the early NBER methods with the modern methods reveals three important differences. First, the A. T. and T. and Snyder indexes that Mitchell seems to have focused on were always analyzed in

detrended form. The modern NBER methods, in contrast, emphasize the use of data that include the secular trend.¹² Second, the rules that the NBER seems to have used to date cycles before 1927 emphasize the choice of the latest possible trough as the reference trough, but make no such adjustment for multiple peaks. This is different from the modern NBER procedures, which choose as turning points both later local peaks and later local troughs. Finally, the component series of the business indexes that Mitchell consulted include many nominal variables. Such variables receive very little weight in modern NBER procedures.

The early NBER's focus on data with the secular trend eliminated could lead to the misclassification of growth recessions as genuine business cycles in the pre-1927 era. Hence, it could cause more cycles to be identified in the early period than in the modern era. The focus on detrended data could also cause systematic differences between early and modern turning points. This can be seen in Figure 1, which shows a stylized business cycle in a series with an upward trend. Points A and B in Figure 1 show the actual peak and trough in the raw data. If the peak and trough are fairly smooth and the upward trend is significant, then the peak in the detrended data will come before the actual peak (at a point such as C) and the trough in the detrended data will come after the actual trough (at a point such as D). Thus, the move from dating cycles using detrended data to dating cycles using data with the trends included could result in the length of early contractions being systematically overstated relative to modern contractions.¹³

The change in the treatment of multiple extremes could also have important implications. In the pre-1927 era, the reference peaks typically coincide with the absolute peaks in either the A. T. and T. or Snyder series, but there is no tendency for them to correspond to the later of the two. Early NBER researchers do, however, seem to have systematically chosen the later of the troughs in the A. T. and T. and Snyder series as the reference date. Because the early NBER procedures move troughs later but

not peaks, they may tend to accentuate the length of contractions and understate the length of expansions relative to postwar procedures that treat multiple peaks and troughs symmetrically.

Finally, differences in the types of series analyzed in deriving the early and modern reference dates may impart some inconsistency to the NBER dates over time.¹⁴ The modern dates are derived almost exclusively from broad measures of real economic activity such as industrial production or employment. The A. T. and T. index and the Snyder clearings index that Mitchell consulted differ from the modern aggregates in that they contain many nominal series. The possible impact of using nominal variables is hard to gauge, but the reliance on nominal series could be an additional source of systematic differences between the early and modern reference dates. If, for example, nominal variables respond more quickly to changes in economic conditions than do real variables, this difference would cause the early NBER business chronology to date both peaks and troughs earlier than the postwar chronology.

III. AN ALGORITHM FOR MATCHING POSTWAR NBER REFERENCE DATES

To test whether these changes in dating techniques have resulted in a meaningful inconsistency in the NBER reference dates over time, in this section I present a simple algorithm that chooses postwar turning points that match the NBER reference dates very closely. This algorithm can be viewed as summarizing the modern NBER procedures.¹⁵ The algorithm can then be used to deduce turning points for the prewar era.

A. Data

Burns and Mitchell do not want to deduce reference cycle dates from any particular aggregate series because such aggregates could move without

the requisite coherence between the various individual series. However, the fear that an aggregate could move without many of the individual components moving is, at least in the postwar era of comprehensive statistical series, largely a theoretical possibility. In practice, series such as real GNP, industrial production, and total employment fluctuate substantially only when many of the individual components fluctuate. As a result, it is useful to see if one can find an algorithm for choosing turning points in some comprehensive aggregate that mimics the NBER postwar reference dates closely. Finding an algorithm that works on a single series is important both because it greatly simplifies the analysis and because there are few comprehensive, high-frequency series that are available for both the postwar and prewar eras.

The aggregate series that I focus on is the Federal Reserve Board's (FRB) index of industrial production. I choose this series for three reasons. First, industrial production is one of the most comprehensive aggregate series that is available monthly. Burns and Mitchell stress that "the monthly reference dates are basic" (1946, p. 80) and therefore, to mimic the NBER procedures, one needs to use a monthly series. Second, the FRB index of industrial production is one of the main series that the current NBER Committee on Business Cycle Dating considers in setting modern reference dates, and it is classified by the NBER as a coincident indicator (see Moore, 1983, p. 7).¹⁶ As a result, it is reasonable to hope that dates derived from this series will match the NBER reference dates closely. Third, in a recent paper, Miron and Romer (1990) derive a new index of industrial production for the period 1884-1940 that can be adjusted to be an adequate prewar extension to the FRB index. As a result, a rule for identifying postwar cycles using industrial production can also be used to deduce cycle dates for the prewar era.

B. Rationale for a Loss Rule

The Burns and Mitchell rules for identifying specific cycles, which

are expressed in terms of duration and amplitude, are complex and cumbersome.¹⁷ Therefore, instead of applying their procedures directly to the postwar FRB index, I derive an alternative dating algorithm that parsimoniously incorporates the duration and amplitude criteria, as well as the other rules for assigning actual dates to turning points.

Burns and Mitchell's idea that contractions must have some minimum duration and amplitude to be classified as business cycles can be well represented by the notion that the amount of output lost during a recession must be of a certain quantity. Figure 2, which shows a stylized picture of the logarithm of industrial production over the course of a recession, illustrates the notion of output loss. If one draws a horizontal line from the peak at point A to point B, the area under this line and above the graph of industrial production shows the cumulative output loss that has occurred between the peak and the time when output returns to its previous peak level.

A rule that says the output loss must be of a particular size for a recession to have occurred should capture very well Burns and Mitchell's notion that both the duration and amplitude of a cycle must be significant for a fluctuation to count as a business cycle. Indeed, the only cases in which this rule might fail are a very short but sharp recession, or a very long but mild one. Fortunately, as a practical matter, such episodes are rare in the United States for the period for which we have industrial production data. Furthermore, in practice the NBER appears to support not the strict application of separate duration and amplitude rules but some balancing of the two. For example, in their justification of the identification of a recession in 1980-1981, Zarnowitz and Moore write: "The 1980 declines in the indicators of major economic activities were relatively short but widespread and deep enough to qualify as another business cycle contraction" (1983, p. 17).

A loss rule can also be used for making concrete Burns and Mitchell's idea that the later of multiple peaks or multiple troughs should be

chosen as the turning point for a series, unless the overall trend of the series has been clearly down between multiple peaks or clearly up between multiple troughs. In terms of output loss, one might say that the rule is to choose the later of two peaks unless the amount of output lost between the first and second peak exceeds some threshold. For troughs, a similar rule could be used, but one would look at the gain in output between the absolute trough and a later local low.

Finally, for extremes that are flat rather than jagged, Burns and Mitchell say that the turning point should be placed at the end of the horizontal stretch. However, they provide little guidance about how much output can move and still be considered part of a flat region. A loss rule, again, can be used to make this criterion concrete: one can impose a rule that says a month is considered part of a plateau if the loss or gain in output between it and the previous month is not greater than some amount.

C. Parameterization of the Loss Rules

The next step is to parameterize the loss rules described above so that when they are applied to data on industrial production, they yield postwar business cycle dates that match the NBER reference dates as closely as possible. For the postwar industrial production series I use the Federal Reserve Board total index. I seasonally adjust the index but do not remove the secular trend, as is consistent with modern NBER procedures.¹⁸ This series is shown in Figure 3.

To deduce the necessary parameters, I first identify every local peak in industrial production and calculate the cumulative loss in log output that occurs between the peak and the date at which industrial production first returns to the peak level.¹⁹ I also identify the actual lowest point in the downturn. Next, I find the cumulative loss between the absolute peak and any later local peaks before the absolute trough in industrial production. I also find the gain between the absolute trough and any later local troughs in the same way. Finally, for every possible peak or trough, I calculate the

loss (or gain) between each of the subsequent three months. Armed with these various measures of loss and gain, I then search for business cycle dating rules that yield cycle dates that match the NBER reference dates as closely as possible.

Cycles. A comparison of the loss in output from each absolute peak until the return to this peak level with the NBER reference dates shows that the smallest postwar downturn that the NBER has classified as a recession is the 1980 cycle. The cumulative loss of log output between the absolute peak in industrial production in March 1980 and the return to this peak level in June 1981 is 0.44; that is, the loss is 44 percentage point-months of industrial production, or about two weeks' output. There are no fluctuations in industrial production with a loss larger than 0.44 that are not counted by the NBER as cycles. Thus, a rule for deciding which fluctuations to count as business cycles that matches the NBER classification exactly is that the cumulative loss of industrial production between the absolute peak and the return to peak must be at least 0.44.²⁰

Multiple Extremes. The loss analysis shows several characteristics of the local peaks and troughs in industrial production that are closest to the NBER reference dates. First, in no case does the NBER reference date correspond to a local peak or trough in industrial production that is before the absolute extreme. Second, one of the most striking features of postwar fluctuations in industrial production is that there is frequently a small dip in output that precedes the main downturn and a small rise that precedes the main upturn. The NBER reference date frequently corresponds to the local peak or trough that is after the absolute high or low in industrial production.

However, the NBER only takes a local peak or trough after the absolute extreme if there has not been a significant rise or fall in industrial production. The largest gain or loss in industrial production between an absolute extreme and a later local extreme that is closer to the NBER reference date is 0.11. This occurs in 1949 when the local trough in October corresponds exactly to the NBER reference date despite the fact that the

absolute trough is in July. There are no cases where the NBER reference date is closer to a later peak or trough than to the absolute extreme and the loss between the absolute extreme and the later extreme is greater than 0.11. There are only two cases where the NBER reference date is closer to the absolute extreme than to a later local extreme despite the fact that the gain or loss between the absolute extreme and the local extreme is less than 0.11.²¹ However, neither of these exceptions to the usual pattern is particularly interesting: the cumulative gain in both instances is very close to 0.11, and the second extreme in each case is so minimal that it is not apparent to the eye from a graph of industrial production.

The consistency of this pattern suggests that the rule for choosing between multiple extremes to match the NBER reference dates is to take the second (or later) of multiple extremes unless the cumulative loss or gain in industrial production between the absolute extreme and the later extreme is greater than 0.11. If the loss or gain is greater than 0.11, then the absolute extreme should be chosen as the turning point.

Plateaus. There is only one case in which the NBER reference date is more than one month later than the peak or trough identified by the multiple extremes rule (though it is occasionally earlier).²² Thus, in deducing the algorithm for matching the NBER dates I only consider whether the first month after a local extreme should be considered a horizontal stretch.

The largest loss or gain in industrial production for which the month after the local extreme consistently corresponds more closely to the NBER reference date is 0.008. Specifically, in every instance where the loss between the local extreme and one month later is less than or equal to 0.008, the later month is closer to the NBER reference date. There are two instances when the month after the local extreme in industrial production is closer to the NBER reference date despite the fact that the loss is greater than 0.008.²³ However, there are enough other cases of losses larger than 0.008 that are not counted that this is the appropriate cutoff point. This

suggests that a rule for dealing with plateaus that will yield dates very close to the NBER reference dates is that the turning point of a cycle should be dated one month after the appropriate local extreme provided that the loss or gain in industrial production is less than 0.008. If the loss or gain is larger than this, then the local extreme itself should be taken as the turning point.

D. Results

When the rules described above are used to identify cycles and deduce turning points from the postwar industrial production series, the resulting dates are very close to the NBER reference dates. This can be seen in Table 2, which shows both the NBER reference dates and the dates chosen by my algorithm.²⁴ Of the 17 reference dates between 1948 and 1990, the dates chosen by the algorithm match the NBER dates exactly in nine instances. The largest discrepancy is five months, which occurs for the NBER trough in 1982. The average difference between the dates derived using the loss algorithm and the NBER dates is 1.1 months.

The differences between the NBER dates and the turning points derived by the algorithm are somewhat systematic. Of the three troughs where the two chronologies differ, the dates derived using the loss rule are all later. The discrepancies for these troughs are also the three largest discrepancies for the 17 postwar turning points. The fact that the algorithm, if anything, tends to date troughs later than does the NBER suggests that the rules may have a slight bias toward accentuating the length of contractions. For peaks there is much less of a pattern. Of the five peaks where the two sets of dates differ, the dates derived by the algorithm are earlier in two cases and later in three. These differences are also all very small.

IV. NEW PREWAR BUSINESS CYCLE DATES

The final step in the evaluation of the consistency of the NBER dates is to apply the dating algorithm to the prewar era. If the criteria for dating cycles have not changed, then the dates that the algorithm chooses for the prewar era should match the NBER dates as closely as do the postwar dates. If not, then this would be evidence that the prewar and postwar NBER reference dates are inconsistent.

This approach rests on two crucial assumptions. First, the algorithm that I derive for the postwar era uses industrial production as the only economic indicator. If this series is a more representative indicator for the postwar economy than for the prewar economy, then the new prewar dates could differ from the NBER dates even if the NBER dates were consistent over time. One piece of evidence that industrial production is roughly as good an indicator for the prewar economy as for the postwar economy is the fact that manufacturing and mining, the two main components of any index of industrial production, have not become a larger or smaller fraction of the economy between 1884 and today. The fraction of national income originating in manufacturing and mining increased from about 21 percent in the late 1800s to 29 percent in the 1930s. After World War II, this fraction reached a high of 31 percent in the 1950s and then declined to about 23 percent in the 1980s.²⁵ Thus, on average, the industrial sector has been equally important in the eras before and after World War II.

The share of output in other sectors of the economy has obviously changed more dramatically over time. However, this change should not affect the usefulness of industrial production as an indicator of aggregate fluctuations. The most important changes, the increasing share of services and the decreasing share of agriculture, have had the effect of substituting one roughly acyclical sector for another. This fact, combined with the relative stability of the size of the industrial sector, suggests that industrial production is a fine indicator of aggregate cyclical activity in both the prewar

and postwar eras.

The second crucial assumption concerns the consistency of the measures of industrial production over time. The postwar algorithm can deduce accurate prewar turning points only if the data to which the algorithm is applied are consistent over time. If there are systematic errors in the prewar measures that are absent from the postwar series, there is no reason to expect the algorithm to date prewar cycles in a way that is consistent with the postwar dates. Therefore, it is important to analyze the consistency of the data.

A. Application of the Loss Rules to the Prewar Era

For the period 1884-1940, Miron and Romer (1990) have created a monthly index of industrial production. This series is not truly consistent with the modern FRB index: it is based on many fewer series than is the modern FRB index, and many sectors of the economy are either over- or under-represented relative to their actual share of value added. The Miron-Romer index is also based very heavily on inputs to the production process and the output of simple manufactures, rather than on the output of highly fabricated products.

However, this index shares many important features with the FRB index. Its most important virtue is that, like the FRB index, it is based on physical production data. The other business indexes for the prewar era include nominal variables such as prices, bank clearings, or interest rates. The other main virtue of the Miron-Romer index is that it has not already been detrended, seasonally adjusted, or otherwise manipulated. This is again in contrast to the existing prewar indexes of industrial production, which are typically only available in highly adjusted forms.

The Federal Reserve Board index of industrial production itself exists on a consistent basis back to 1923. For the four years 1919-1922 there is also official data, though the FRB has not attempted to make these observations strictly comparable to the postwar series.²⁶ Since the Miron-

Romer index and the FRB index overlap for the period 1919-1940, it is possible to see just how similar these two indexes actually are. Because the Miron-Romer index is based largely on inputs and simple manufactures, it is more volatile than the FRB index. In addition, because it is constructed using relatively few commodities, it tends to have more random movements.

Given these obvious differences between the two series, it would be inappropriate to simply apply the dating algorithm derived from the postwar FRB index to the prewar Miron-Romer index. Instead, I first adjust the Miron-Romer index to be more comparable to the FRB index. To do this, I use a regression to estimate the relationship between the FRB index and the Miron-Romer series in a period of overlap. I then use this estimated relationship to form adjusted values for the Miron-Romer index for the period before 1919. The details of this adjustment are given in the Appendix. The resulting prewar index of industrial production that combines the FRB index after 1919 and the adjusted Miron-Romer series through 1918 is given in Figure 4.

The application of the postwar rules given Section III to the prewar index of industrial production is straightforward. Because the adjusted prewar index only extends back to 1884, I am only able to identify cycles beginning in the late 1800s. In applying the algorithm to the prewar era, I carry the analysis through 1940. The dates that the postwar algorithm chooses as turning points for prewar business cycles are shown in Table 3. Table 3 also gives the corresponding NBER reference dates.

B. Comparison of the New Dates and the NBER Reference Dates

Cycles. Table 3 shows several similarities and differences between the dates chosen by the postwar algorithm and the prewar NBER reference dates. The strongest similarity involves what counts as a cycle. There is only one fluctuation between 1884 and 1940 that the NBER identifies as a cycle that the postwar algorithm does not choose. This is the fluctuation of 1890-1891. This cycle is not counted by the postwar algorithm because the

total output loss between the absolute peak and the return to peak is only 0.25, while the smallest loss postwar that counts as a cycle is 0.44. This cycle is one that other researchers have frequently mentioned as being questionable.²⁷ Thus, its exclusion will probably surprise few. The fact that there are not more instances of slowdowns being classed as cycles may indicate that Mitchell set the criterion for what counts as a cycle in detrended data strict enough that most cycles correspond to a significant fall in real output. The fact that the Business Annals played a key role in the identification (but not the actually dating) of cycles could also explain why there are not more questionable prewar cycles in the NBER chronology. It is likely that the business press only took notice of significant declines in production.

Several other prewar fluctuations that others have questioned are confirmed by the application of the postwar algorithm. For example, the cycles of 1887-1888 and 1926-1927 are often alleged to have been mere slowdowns in growth rather than actual declines in real output (see, for example, Ayres, 1939, and Moore and Zarnowitz, 1986). However, the output loss in both instances (0.58 in 1887-1888 and 0.68 in 1926-1927) is sufficient for them to be counted as cycles according to postwar NBER criteria. Similarly, the immediate post-World War I recession in 1918-1919 that is often skipped by other chronologies also appears to be a genuine cycle according to modern criteria.

There are two fluctuations in the prewar era that the postwar algorithm classifies as cycles but the NBER does not. These cycles occur in 1916-1917 and 1939-1940. Both these fluctuations are reasonably small (the loss in 1916-1917 is 0.46 and that in 1939-1940 is 0.65) and brief, and both are associated with the start of war in Europe. Because of the special circumstances associated with these fluctuations, their inclusion does not constitute a major discrepancy with the NBER. Overall, the close correspondence between what the NBER classifies as prewar cycles and what the postwar algorithm identifies as cycles suggests that the NBER has

been quite consistent over time in the identification of cycles.

Turning Points. There is much less similarity between the dates of peaks and troughs in the NBER chronology and the new chronology derived from the algorithm. Of the 16 turning points between 1887 and 1917 for which both chronologies date a cycle, there is exact agreement on the date of the peak or trough in only five instances. The average absolute value of the discrepancy between the NBER dates and the new dates for this period is 4.5 months. The largest discrepancy occurs for the peak immediately before World War I: the NBER dates the peak in January 1913 while the postwar algorithm chooses June 1914. There is more agreement for the turning points in the two decades between the world wars. Of the 12 turning points between 1918 and 1940 for which both chronologies date a cycle, there is exact agreement on the month and year of the peak or trough in six instances, and the average discrepancy is 1.6 months.²⁸

Not only is there little agreement between the turning points in the two chronologies, but the differences appear to be systematic. Of the nine peaks between 1887 and 1940 for which the two chronologies do not agree, the new date that I derive is later than the NBER peak in seven instances. Of the eight troughs between 1887 and 1940 for which the two chronologies do not agree, the new date that I derive is earlier than the NBER trough in seven instances. Averaged over all 14 cycles that are identified in both chronologies, the new peaks lag the NBER peaks by an average of 3.3 months, and the new troughs lead the NBER troughs by an average of 2.8 months. Clearly, the new dates that I derive systematically date peaks later and troughs earlier than does the NBER prewar business cycle chronology.

C. Tests of the New Dates

That the new dates I derive using the postwar algorithm are significantly and systematically different from the prewar NBER turning points provides statistical evidence that there is an important inconsistency between the early and modern NBER reference dates. The obvious

weakness in this statistical evidence is the possibility that the index of industrial production to which the algorithm is applied is fundamentally different between the prewar and postwar eras. To check for this possibility, I conduct three tests of the reliability of the new dates.

First, the adjusted Miron-Romer index and the FRB index both exist for the period 1919-1940. In the derivation of the new business cycle dates given in Table 3, I apply the postwar algorithm to the adjusted Miron-Romer index only for the dates before 1919; for the dates in the period 1919-1940 I use the FRB index. To see if the adjusted Miron-Romer series is genuinely similar to the FRB index, I also apply the dating algorithm to the adjusted Miron-Romer series for 1919-1940. The dates derived from the adjusted Miron-Romer index for this era are quite similar to those derived from the FRB index. Both series identify the same cycles. There is a slight tendency for the dates based on the Miron-Romer series to lag those based on the FRB index (the average turning point is 2.3 months later), but because peaks and troughs lag equally, there is no systematic difference in the length of contractions and expansions.

Second, there exists an unofficial continuation of the FRB index back to 1899 constructed by Woodlief Thomas that can be used to date prewar cycles.²⁹ This series has some obvious weaknesses; most notably, it only exists for a limited period and the specifics of its derivation are not described anywhere. However, it does appear to be based on many of the same series as the later FRB index and thus is likely to be reasonably consistent with the FRB series.

Applying the postwar algorithm to the Thomas index for 1899-1919 yields turning points that are very similar to those derived from the adjusted Miron-Romer index. Except for the cycle just before World War I, the two series identify the same cycles. Most of the dates are within a month or two of each other.³⁰ As with the FRB index for 1919-1940, the dates derived from the Miron-Romer index for both peaks and troughs appear to be a few months later than those in the Thomas index; the average lag is 2.2 months.

As before, however, the whole cycle is moved a few months later; thus the lengths of expansions and contractions are essentially unaffected. The fact that the postwar algorithm yields similar dates when applied to various reasonable prewar indexes of industrial production suggests that differences between the new dates and the NBER reference dates reflect changes in NBER methods, not changes in the data used to identify cycles.

A final test of the robustness of the new dates is to apply the postwar algorithm to a less adjusted version of the prewar Miron-Romer index. The adjustments to the Miron-Romer series described in the appendix not only smooth and seasonally adjust the raw series, but also dampen cyclical fluctuations. If one forces the adjustments to leave cyclical fluctuations as severe as in the raw series, the dates derived by applying the algorithm to the less adjusted index are very similar to those given in Table 3.³¹ Most importantly, the dates derived from the undamped prewar series still place peaks later and troughs earlier than does the NBER chronology. This suggests that the new dates are robust to even unreasonably extreme variations in the underlying data.

V. IMPLICATIONS OF THE NEW BUSINESS CYCLE DATES

A. Duration

Because the new prewar dates that I derive differ systematically from the NBER reference dates, the more consistent business cycle chronology yields very different conclusions about changes in business cycles over time. Most obviously, the new dates radically alter one's view of changes in the duration of contractions and expansions over time. Table 4 shows the average duration of contractions and expansions in various time periods using both the NBER reference dates and the new dates. The periods that I consider are 1887-1917, 1918-1940, and 1948-1990.

As has been noted in many previous studies, using the NBER

reference dates for all periods leads to the conclusion that there has been a significant decline in the length of contractions and a tremendous rise in the length of expansions over time.³² For both the pre-World War I period (1887-1917) and the interwar period (1918-1940), the average length of contractions using the NBER dates is approximately 18 months, while in the period after 1948, the average length of contractions is just 11 months. For expansions, the average duration is roughly half as long for the two periods before World War II as for the postwar era.

The results are very different when one considers the new dates. For the postwar era I report estimates of duration derived from both the NBER reference dates and the dates that result from the application of my algorithm. Though these two postwar chronologies are very similar by construction, the duration measures indicate that the turning points chosen by the algorithm yield postwar contractions that are on average about one month longer than NBER contractions. Because this tendency of the algorithm to accentuate slightly the length of contractions will be present in all periods, the appropriate comparison is between the dates chosen by the algorithm in each era.

When one compares my new dates for 1887-1917 to the postwar dates chosen by the algorithm, there is no evidence of a decline in the duration of contractions over time. On the contrary, the average contraction appears to increase by nearly three months between the pre-World War I and the post-World War II eras. This same result holds even more strongly when one considers the median rather than the mean contraction duration: the median contraction using the new business cycle dates lasts eight months in the pre-World War I era and 12 months in the postwar period.³³

The new business cycle dates yield two other duration findings that are more plausible than the results based on the NBER dates. First, using the NBER chronology, the average duration of contractions is almost identical in the pre-World War I and interwar periods. Using the new dates, however, contractions in the interwar period are roughly six months longer

on average than in the pre-World War I era. This finding is much more consistent with the usual perception that the Great Depression, which dominates the interwar era, is an unusually long contraction. Second, using the NBER dates for the pre-World War I and interwar eras there is a suspicious similarity in the length of expansions and contractions. This is exactly what one would expect to find if the early NBER dates were derived from detrended data. There is much less similarity in the length of contractions and expansions using the new business cycle chronology. The dates chosen by the algorithm for the period before 1940 show contractions to be only about a third as long as expansions.

Even using the new dates, there is evidence of an increase in the average length of expansions over time. The average expansion in the period 1887-1917 using the new dates is 32 months, while the average postwar expansion is 50 months. This is less of a change than is shown by the NBER dates, but it is still quite large. However, much of the long average duration of postwar expansions is due to the very long expansion in the 1960s. If one considers the median rather than the mean expansion, there is much less lengthening of expansions over time: the median expansion is 31 months in the period 1887-1917 and 41 months in the period 1948-1990.

B. Severity

The loss measures calculated in the derivation of the new dates also change one's view of the severity of the typical recession in different eras. Table 5 shows the loss in industrial production from the absolute peak to the return to this peak for each recession in the periods 1887-1917, 1918-1940, and 1948-1990.³⁴ The recessions are ordered according to severity.

Based on these measures, the mean and median loss in industrial output in recessions is very similar before and after the interwar era. For the period 1887-1917, the mean loss associated with a recession is 1.36 and the median loss is 1.16. For the period 1948-1990, the mean loss is 1.29 and

the median loss is 1.19. This suggests that recessions have not, on average, become less severe over the twentieth century. In the interwar era, however, the typical loss is much higher: the mean loss associated with a recession is 6.79 and the median loss is 1.89. The difference between the mean and the median for this period is dramatic testimony to the severity of the Great Depression.

The similarity of the typical loss associated with a recession in the pre-World War I and post-World War II eras masks an important change in the distribution of the severity of recessions over time. In both the pre-1917 era and the interwar period there is a wide range of cycles, including many mild cycles and many severe cycles. In the postwar era, in contrast, most cycles fall in the moderate range. This change suggests that the distribution of cyclical severity has primarily narrowed over time, rather than shifted uniformly toward milder cycles.

This change in the distribution of cycles may affect how one views the apparent lengthening of expansions over time. Some of the greater length of postwar expansions comes from the fact that there are many mild cycles in the pre-World War I and interwar eras. In the postwar era the second smallest recession (1960-1961) involves an output loss of 0.96. In the pre-1917 period, in contrast, there are four recessions with output losses between 0.44 and 0.96, and in the interwar era there are three such contractions. The greater frequency of mild cycles causes expansions to look shorter in the pre-World War I and interwar eras than in the post-World War II era, but nearly all of this change would disappear under a more stringent definition of what counts as a cycle.³⁵

C. Persistence

Another factor to consider when analyzing the length of expansions is the time that it takes to recover to the previous peak level of output. Sichel (1992) describes this time between the trough and the return to peak as a third stage of the business cycle and argues that it indicates how rapidly

the effects of a recession are undone. The new dates, in conjunction with the adjusted prewar index of industrial production, show that in the pre-World War I era it took the U.S. economy an average of 8.9 months to move from the trough to the previous peak.³⁶ In the postwar era the average time to recovery is 10.0 months, and in the interwar era it is 17.0 months.

The largest difference is obviously between the interwar era and both the pre-World War I and post-World War II periods. Some of this greater interwar time to recovery reflects the severity and persistence of the Great Depression. However, even the median time to recovery is substantially longer in the period 1918-1940 than in the decades before and after, because the recoveries from the recessions of 1921, 1924, and 1938 are all relatively slow.

The difference in the time to recovery between the pre-World War I and the post-World War II periods is small -- just 1.1 months. However, it does indicate that the economy was perhaps slightly more resilient before World War I than it is today. This shorter prewar time to recovery compounds the earlier finding that the average time from peak to trough is nearly three months shorter before 1917 than after 1948. These two findings taken together imply that the average time that the economy spends below the previous peak is about a third of a year shorter before World War I than after World War II. Thus, while the average loss in output is the same in the two eras, the loss is more concentrated in the pre-World War I era than today.

VI. CONCLUSION

This paper provides two types of evidence of inconsistency in the NBER chronology of business cycle peaks and troughs. First, an analysis of the methods used to derive the early and modern reference dates shows

that the NBER procedures for dating peaks and troughs have changed substantially over time. Most importantly, there was much more emphasis in the pre-1927 era on using detrended data and on dating troughs, but not peaks, as late as possible. Second, statistical analysis shows that an algorithm that deduces postwar turning points that are nearly identical to the NBER reference dates yields prewar turning points that are systematically different from those of the NBER. The new chronology that I derive consistently dates prewar peaks later and prewar troughs earlier than does the NBER.

Removing the inconsistencies in the NBER reference dates invalidates the usual view that recessions have gotten shorter over time. Using the new business cycle chronology that I develop, the length of contractions actually increases by roughly three months between the pre-World War I and post-World War II eras. At the same time, even the consistent business cycle dates show that expansions have become longer over time.

The new dates and the revised index of industrial production from which they are derived also show that the average loss in output is nearly identical in pre-World War I and post-World War II contractions. However, there has been a compression of the distribution of the severity of cycles: in the pre-World War I and interwar eras there are both more mild cycles and more severe cycles than in the post-World War II period. The average time that it takes industrial production to recover from its trough to the previous peak level has also increased slightly over time, suggesting that cycles are somewhat more persistent today than they were in the past.

Thus, the changes in recessions revealed by the new chronology do not show an obvious improvement in business cycles over time. While the time separating contractions has become longer between the prewar and postwar eras, recessions themselves have not on average become shorter, less severe, or less persistent over time. Inconsistencies in the way that the NBER reference dates have been set led researchers to mistake mere

changes in dating procedures for genuine changes improvements in economic stability.

APPENDIX

To adjust the Miron-Romer index of industrial production for 1884-1918 to be more consistent with the modern FRB index, I run a regression between the two series over the period 1923-1928. I choose this limited period rather than the full period of overlap for two reasons. First, it is undesirable to use the period 1919-1922 because the FRB index for this period may be inconsistent with later FRB data. Therefore, a regression based on this period might not yield a prewar series consistent with the postwar series. Second, it seems wise to exclude the boom of 1929 and the Great Depression because the FRB index may behave unusually in extreme times.³⁷ The 1923-1928 period does, however, include two recessions, so there is ample variation from which to estimate a relationship.³⁸

The specification that I use regresses the log level of the FRB index (which is not seasonally adjusted) on a constant, a trend, eleven monthly dummy variables, the contemporaneous log level of the Miron-Romer index, and six lags and six leads of the Miron-Romer index.³⁹ The contemporaneous value of the Miron-Romer series is obviously included to capture the main relationship of interest. The constant and the monthly dummies are present to take into account seasonal fluctuations. The trend takes into account possible differences in the trend of the two series over the mid-1920s.⁴⁰ The six lags and leads of the Miron-Romer index are included to allow for the possibility that the timing of the two indexes could be different at a fairly short horizon.

The results of this regression suggest that there is a very close relationship between the two industrial production series. The R^2 of the regression is 0.90. The sum of the coefficients on the lags and leads of the Miron-Romer index is 0.67 with a standard error of 0.10. The contemporaneous value of the Miron-Romer series, the first lead, and the second lag are the largest and most significant individual coefficients. There does not appear to be any systematic difference in timing between the FRB

and the Miron-Romer indexes.

To form the adjusted Miron-Romer index for the period before World War I, I first regress the Miron-Romer index for 1884-1918 on a constant, a trend, and eleven monthly dummy variables and form a seasonally adjusted series by removing the effect of the monthly dummy variables. I then use the estimated coefficients from the regression for the 1920s to combine the lags and leads of this index. Because the seasonal effects are removed in a separate step, I do not use the seasonal coefficients in forming these fitted values.⁴¹ This procedure allows for the possibility that seasonal movements may have changed between the turn of the century and the 1920s. The final prewar index of industrial production that I use merges the adjusted Miron-Romer series for 1884-1918 with the FRB index for 1919-1940.⁴² By construction, the series match up very closely in 1919.

REFERENCES

- Axe, Emerson Wirt and Houghton, Ruth, "Financial and Business Cycles, Manufacturing Growth, and Analysis of Individual Industries, 1883-1930," The Annalist, January 16, 1931, 37, 94-98 and 150-151.
- Ayres, Leonard P., Turning Points in Business Cycles, New York: Macmillan, 1939.
- Bry, Gerhard and Boschan, Charlotte, Cyclical Analysis of Time Series: Selected Procedures and Computer Programs, New York: NBER, 1971.
- Burns, Arthur F., ed., Wesley Clair Mitchell: The Economic Scientist, New York: NBER, 1952.
- Burns, Arthur F. and Mitchell, Wesley C., Measuring Business Cycles, New York: NBER, 1946.
- Cloos, George W., "How Good Are the National Bureau's Reference Dates?" Journal of Business, January 1963, 36, 14-32. (a)
- _____, "More on Reference Dates and Leading Indicators," Journal of Business, July 1963, 36, 352-364. (b)
- Diebold, Francis X. and Rudebusch, Glenn D., "Have Postwar Economic Fluctuations Been Stabilized?" American Economic Review, 1992, forthcoming.
- Fels, Rendig, American Business Cycles, 1865-1897, Chapel Hill: University of North Carolina Press, 1959.
- Miron, Jeffrey A. and Romer, Christina D., "A New Monthly Index of Industrial Production, 1884-1940," Journal of Economic History, June 1990, 50, 321-332.
- Mitchell, Wesley C., Business Cycles, Memoirs of the University of California, Vol. 3, Berkeley: University of California Press, 1913.
- _____, Business Cycles: The Problem and Its Setting, New York: NBER, 1927.
- _____, "Testing Business Cycles," News-Bulletin of the NBER, March 1, 1929, No. 31. (a)

- _____, "A Review," in Recent Economic Changes in the United States, Report of the Committee on Recent Economic Changes, of the President's Conference on Unemployment, Vol. II, New York: McGraw-Hill for NBER, 1929. (b)
- _____, "Business Cycles," in Encyclopedia of the Social Sciences, Vol. 3, New York: Macmillan, 1930, pp. 92-107.
- _____, What Happens During Business Cycles? A Progress Report, New York: NBER, 1951.
- Mitchell, Wesley C., assisted by Arthur F. Burns, A Preliminary Draft of Chapters I, II, and III of Business Cycles, Volume II, Analysis of Cyclical Behavior, unpublished manuscript, circa 1937.
- Moore, Geoffrey H., "What is a Recession?" in Geoffrey H. Moore, ed., Business Cycles, Inflation, and Forecasting, Cambridge, Ma.: Ballinger Publishing for NBER, 1983, pp. 3-9.
- Moore, Geoffrey H. and Zarnowitz, Victor, "The Development and Role of the National Bureau of Economic Research's Business Cycle Chronologies," in Robert J. Gordon, ed., The American Business Cycle, Chicago: University of Chicago Press for NBER, 1986, pp. 735-779.
- National Bureau of Economic Research, News-Bulletin, various issues.
- Persons, Warren M., Forecasting Business Cycles, New York: John Wiley & Sons, 1931.
- Romer, Christina D., "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869-1908," Journal of Political Economy, February 1989, 97, 1-37.
- Rorty, M. C., "The Statistical Control of Business Activities," Harvard Business Review, January 1923, 1, 154-166.
- Sichel, Daniel, "Inventories and the Three Phases of the Business Cycle," Working paper, Board of Governors of the Federal Reserve System, August 1992.
- Snyder, Carl, Business Cycles and Business Measurements, New York: Macmillan, 1927.
- Thorp, Willard Long, Business Annals, with an introductory chapter by Wesley C. Mitchell, New York: NBER, 1926.

- Trueblood, Lorman C., "The Dating of Postwar Business Cycles," Proceedings of the Business and Economic Statistics Section of the American Statistical Association, 1961.
- U.S. Board of Governors of the Federal Reserve System, Federal Reserve Bulletin, January 1941.
- U.S. Board of Governors of the Federal Reserve System, "New Federal Reserve Index of Industrial Production," Federal Reserve Bulletin, August 1940, 26, 753-769.
- U.S. Bureau of Economic Analysis, The National Income and Product Accounts of the United States, 1929-1982, 1986.
- U.S. Bureau of the Census, Historical Statistics of the United States, 1975.
- Watson, Mark W., "Business Cycle Durations and Postwar Stabilization of the U.S. Economy," NBER Working Paper No. 4005, March 1992.
- Zarnowitz, Victor, "On the Dating of Business Cycles," Journal of Business, April 1963, 36, 179-199. (a)
- _____, "Cloos on Reference Dates and Leading Indicators: A Comment," Journal of Business, October 1963, 36, 461-463. (b)
- _____, "Business Cycles and Growth: Some Reflections and Measures," in W. J. Muckl and A. E. Ott, eds., Wirtschaftstheorie und Wirtschaftspolitik: Gedenschrift fur Erich Preiser, Passau: Passavia Universitatsverlag, 1981, pp. 475-508.
- Zarnowitz, Victor and Moore, Geoffrey H., "The Timing and Severity of the 1980 Recession," in Geoffrey H. Moore, ed., Business Cycles, Inflation, and Forecasting, Cambridge, Ma.: Ballinger Publishing for NBER, 1983, pp. 11-17.

NOTES

1. See, for example, Moore and Zarnowitz (1986) and Diebold and Rudebusch (1991).
2. Moore emphasizes the importance of this point when he writes: "Nor is a mere slowing down or cessation of growth in total activity enough to qualify as a contraction" [1983, p. 4].
3. Bry and Boschan (1971, p. 27) suggest that Burns and Mitchell may also have an implicit rule that the expansion and contraction phases of a specific cycle must each be at least five months long.
4. According to Mitchell (1951, p. 12) the reference dates for the period 1919 to 1938 were revised slightly after the publication of Measuring Business Cycles. He does not, however, discuss how or why these revisions were made.
5. This article mentions the fact that Simon Kuznets supervised the empirical work for the study. This is useful because it may explain why in Measuring Business Cycles Burns and Mitchell say that "Simon Kuznets took a leading part in the preparation of the original set of reference dates" (1946, p. 80). It seems clear that Kuznets's role was as a junior collaborator working under Mitchell on this project.
6. This information is from the News-Bulletin of the NBER for various years in the 1930s.
7. Mitchell (1913) also compiled dates for peaks and troughs from contemporary business publications. These dates, however, do not appear to have been used explicitly in the setting of the NBER reference dates.
8. To actually deduce dates from Chart VI of Business Annals I assume that the peak occurs immediately before the onset of recession and that the trough occurs immediately before the onset of revival. This interpretation fits with the fact that in "Testing Business Cycles" (1929) Mitchell dates recessions as beginning the month after the NBER peak and revivals as beginning the month after the NBER trough. I only try to read the dates on the chart down to the quarter. Though Mitchell says that the chart "shows for every cycle the quarters and years which Dr. Thorp has taken as marking off recession and revival" (1926, p. 64), it is clear that the chart is marked off in months. However, the chart is sufficiently crude that it would be difficult to read off the dates at the monthly level with much accuracy. In the

comparison with the NBER reference dates, I use the NBER dates denominated in quarters from Moore and Zarnowitz (1986, p. 750).

9. See Mitchell (1927, pp. 294-295 and 304-305) for a description of the two series. The A. T. and T. index is available in Rorty (1923, p. 159). The Snyder index is available in Snyder (1927, pp. 292-293).

10. The NBER reference peak in 1918 is set much later than the crest in either of the series, presumably because of the war.

11. Numerous other researchers also dated business cycles in the 1930s (see, for example, Persons, 1931, Axe and Houghton, 1931, and Ayres, 1939). Without exception, these authors focused on detrended aggregate series similar to those used by Mitchell. Thus, the correlation between the early NBER reference dates and these other chronologies that Burns and Mitchell note (1946, p. 108) should not be surprising.

12. Mitchell's writings over the period 1927-1946 show a definite change in his views toward detrending that parallels the apparent change in methods. Business Cycles (1927, pp. 212-233) contains a careful discussion of methods of detrending and specifically states that because "cyclical fluctuations stand out more clearly after the statistical trends and seasonal variations have been eliminated ..., we shall therefore make such use as we can of these eliminations in our further work" (Mitchell, 1927, pp. 258-259). This same view is expressed again three years later in Mitchell (1930, p. 94). In a preliminary draft of Measuring Business Cycles written by Mitchell alone, he expresses ambivalence about detrending. In response to the rhetorical question, "Ought we not eliminate secular trends from the original data of our time series before taking cyclical measurements?" Mitchell responds that there "would be substantial gains" but "they would be attended by a decrease in the significance of the results for our ultimate purpose" (1937, Chap. III, pp. 48-49). This debate is dropped from the final version of Measuring Business Cycles, which simply states "we take as our basic unit of analysis a business cycle that includes the portion of secular trend falling within its boundaries" (1946, p. 270).

13. Zarnowitz (1981, p. 506) contains a long discussion of the fact that the identification of prewar NBER reference dates "relied to a considerable extent on business annals and trend-adjusted indexes of business conditions". While Zarnowitz realizes that the use of detrended data could lead to the misclassification of growth recessions as business cycles, he does not consider the possibility that it could also result in the systematic misdating of cycles. Moore, on the other hand, realizes that the prewar reference dates could be systematically biased. Trueblood (1961, p. 16) writes that "a recent statement by Geoffrey Moore in an unpublished National Bureau staff report found that 'the National Bureau business cycle reference dates between 1885

and 1914 (possibly earlier) seem to be biased in the sense that peaks are frequently predated and troughs are frequently postdated." Trueblood reports that Moore believes "a rough correction of this bias has the effect of lengthening the mean duration of expansions, 1854-1914, by 4 months and of decreasing that of contractions by an equal amount" (1961, p. 16). Moore, however, in all of his subsequent work on the duration of cycles has taken the NBER reference dates as given and has never published a major revision of the prewar reference dates.

14. Watson (1992) also considers possible inconsistencies in the NBER reference dates. He shows that if the early dates were based on many volatile series while the postwar dates are based on smoother aggregate series, this could cause there to be more cycles and for peaks and troughs to be systematically misdated in the prewar era. However, my research suggests that early dates were not based on the many volatile series that Measuring Business Cycles analyzes and describes. Rather, they are based on only a few aggregate series. Furthermore, the A. T. and T. and Snyder series that Mitchell appears to have used to date cycles are not significantly more volatile than the adjusted Miron-Romer-FRB series that I use. (For example, the standard deviation of the deviations from trend of the Miron-Romer-FRB series for 1884-1921 is 7.70 while that for the Snyder series for the same period is 8.95 and that for the A. T. and T. series is 9.34.) Hence while this hypothetical bias could explain the inconsistency I find in Section IV, it does not appear to be the actual source of the misdating of prewar cycles.

15. In this exercise, I take the postwar NBER reference dates as given. However, the postwar dates have been the subject of a vigorous debate involving Cloos (1963a and 1963b), Trueblood (1961), and Zarnowitz (1963a and 1963b).

16. Cloos (1963a) also advocates the use of the FRB index for setting reference dates.

17. Bry and Boschan [1971] develop a computer algorithm for mimicking NBER specific cycle dating procedures that is indeed complicated and involves various ad hoc adjustments.

18. I use the revised version of the index which is on a 1986 base and is available without seasonal adjustment from the Federal Reserve Board. To remove seasonal fluctuations, I regress the logarithm of the unadjusted series on a constant, eleven monthly dummy variables and a linear trend. I then use the residuals, with the trend and constant added back in, as the seasonally adjusted series. I seasonally adjust the postwar series myself because I want to be able to adjust the prewar Miron-Romer series in the same way.

19. I start these calculations in 1948 because World War II causes such a spike in industrial production that a loss rule would be forced to say that there was one long cycle from December 1944 until the return to that peak level in December 1950.

20. The largest postwar fluctuation that is not counted by the NBER as a reference cycle occurs in the middle of 1959 and has a cumulative loss of 0.34. Thus, a grey area for identifying cycles exists between a loss of 0.34 and 0.44: it is not possible to deduce what the NBER would do with a cycle with a loss of, say, 0.40. For all of my rules I take the loss or gain that the NBER has counted in the postwar era as the cutoff point. However, I have also tried using the alternative rules based on the loss or gain that the NBER has not counted in the postwar era. When these weaker rules are applied to the prewar era some of the individual prewar dates change, but the changes are small enough and rare enough that they do not change any of the conclusions of the paper.

21. One occurs in 1975, when the absolute trough in March rather than a local trough in July corresponds to the NBER reference date. The other occurs in 1982, when the local trough in October 1982 rather than the later low in March 1983 is closer to the NBER trough.

22. This occurs in 1969 when the peak in industrial production is in October and the NBER peak is in December. The fall in industrial production in this case, however, is quite large between October and December, suggesting that vagaries in the industrial production series, rather than a plateau rule, explain this exception.

23. These occur in 1948 (where the loss between October and November is 0.015) and in 1969 (where the loss between October and November is 0.010).

24. The NBER reference dates for all periods are from Moore and Zarnowitz (1986, p. 750).

25. The data for the period before 1969 are from Historical Statistics (1975, p. 238) The data for the period after 1969 are from the National Income and Product Accounts (1986, p. 256).

26. The interwar FRB index is described in U.S. Board of Governors (1940).

27. Zarnowitz, for example, lists it among the mildest prewar cycles (1981, p. 503), and Fels describes it as "singularly mild" (1959, p. 159).

28. The fact that there is more similarity between the dates chosen by the algorithm and the NBER peaks and troughs in the interwar era than in the pre-World War I period is not surprising for two reasons. First, the change in procedures identified in Section II occurred in 1927. Second, while the reference dates before 1918 have never been revised, some of the reference dates for the interwar period were changed slightly in the early 1950s. These revisions were so small that it does not appear that there was a thorough revamping of the interwar dates. However, of the five dates that were changed (the original dates were 1919:4, 1921:9, 1927:12, 1929:6, and 1938:5), four of the new dates are closer to the dates chosen by the algorithm than were the original NBER turning points. Because the revisions conducted by researchers familiar with postwar dating methods move the interwar NBER dates in the direction indicated by my algorithm, there is less discrepancy between the new dates and the NBER dates for the interwar era.

29. This series is given in Ayres (1939, p. 202-203). Annual values are published in the Federal Reserve Bulletin (1931, p. 46).

30. In nearly every case the date derived from the Thomas index using the postwar algorithm is much closer to the date derived from the Miron-Romer series than to the NBER reference date. For example, while the NBER reference peak is September 1902, the peak derived from the Miron-Romer series is July 1903, and that from the Thomas series is June 1903.

31. This less adjusted version of the Miron-Romer series is derived by imposing the constraint that the sum of the coefficients on the leads and lags of the Miron-Romer series in the regression described in the appendix is equal to one. It is important to smooth the Miron-Romer index because the series has several isolated extremes that are due to temporary industry-specific disturbances such as strikes or tariffs.

32. See, for example, Moore and Zarnowitz (1986), and Diebold and Rudebusch (1992).

33. Two experiments can be used to identify the contribution of the various changes in NBER procedures to the difference between the average length of contractions as measured by the early NBER chronology and as measured by the new dates. First, to quantify the importance of the use of detrended data, I compare the absolute peaks and troughs in the log level of the prewar industrial production series with the peaks and troughs in detrended industrial production. Because many peaks and troughs are sharp rather than smooth, the turning points in the two series are often the same. But in several key cases (most notably the peaks in 1913-1914 and 1926-1927), the turning points in the detrended series are much closer to the NBER reference dates than are the turning points in the raw series. A comparison of the average duration of contractions in the raw data and in the detrended

industrial production series implies that about half of the difference in the length of contractions between the early NBER reference dates and the dates produced by the algorithm is due to the NBER's use of detrended data.

Second, to measure the importance of the early NBER's procedure of using the later of multiple troughs, but not of multiple peaks, I consider the effects of choosing a later trough rather than the absolute trough in detrended industrial production if there is a plausible candidate and the NBER chose a late trough. Again, this change has a large effect in several important cases (most notably 1887-1888 and 1911-1912). The experiment implies that the NBER's policy of adjusting troughs but not peaks accounts for approximately one third of the difference in the average length of early contractions as measured by the NBER and by the algorithm. The remainder of the difference is most likely due to the early NBER's use of nominal variables, the application of subjective judgement by the NBER, and random factors.

34. I use this measure of severity rather than the cumulative loss of industrial production from the peak chosen by the algorithm to the return to the algorithm peak, because it is the measure used to initially decide whether a downturn is severe enough to count as a recession. The results are very similar, however, if the alternative measure of severity is used.

35. For example, eliminating all cycles with a loss of less than 0.85 from the new chronology would result in an average expansion of 52.6 months in the pre-World War I era and 58.0 months in the post-World War II era.

36. This calculation is based on the peaks and troughs chosen by the algorithm. The results are very similar if absolute peaks and troughs are used.

37. Romer (1989) shows that even series that are traditionally quite stable become very volatile during the Great Depression.

38. I have tried using other sample periods, including the full sample 1919-1940. The results are quite robust.

39. This procedure is equivalent to seasonally adjusting both the FRB and the Miron-Romer series and then regressing the seasonally adjusted series on one another.

40. Instead of including a trend in the regression, one could simply run the regression in differences. The results of this regression are nearly identical to those reported. I choose to work in levels because it facilitates filtering the prewar observations.

41. I do, however, add in the constant and the trend from the 1920s regression. This essentially adjusts the pre-1918 trend by an amount equal to the usual difference between the trend of the FRB index and the trend of the Miron-Romer index in the period 1923-1928.

42. I use the FRB index beginning in 1919 even though it is not strictly comparable to the later index because it has significantly broader coverage than the Miron-Romer index. Thus, while the FRB index for 1919-1922 may have some systematic errors that would be eliminated by using the adjusted Miron-Romer index, it also potentially has additional information that would be lost using the less broad index.

TABLE 1
Peaks and Troughs in Business Indexes

<u>NBER Reference Dates</u>		<u>A. T. and T. Index</u>		<u>Snyder Clearings Index</u>	
<u>Peak</u>	<u>Trough</u>	<u>Peak</u>	<u>Trough</u>	<u>Peak</u>	<u>Trough</u>
1887:3	1888:4	1887:3	1888:3	1887:6	1888:3
1890:7	1891:5	1890:5,10	1891:5	1890:7	1891:3
1893:1	1894:6	1892:2	1894:6	1893:1	1893:10
1895:12	1897:6	1895:10	1896:10	1895:12	1897:3
1899:6	1900:12	1899:12 1900:2	1900:11,12	1899:6	1900:9
1902:9	1904:8	1902:9	1903:12	1901:6	1904:7
1907:5	1908:6	1907:5,7	1908:5,6	1906:2	1908:1
1910:1	1912:1	1910:1,3	1911:4	1910:4	1911:12
1913:1	1914:12	1913:1	1914:12	1913:2	1914:12
1918:8	1919:3	1916:11 1917:1	1919:3	1916:12 1917:1	1919:3
1920:1	1921:7	1920:1	1921:4,5,7	1919:8,9	1921:3
1923:5	1924:7	1923:5	1924:6	1923:5	

Note: The dates are expressed year:month.

TABLE 2
Postwar Peaks and Troughs

<u>NBER Reference Dates</u>		<u>Dates Produced by the Algorithm</u>	
<u>Peak</u>	<u>Trough</u>	<u>Peak</u>	<u>Trough</u>
1948:11	1949:10	1948:10	1949:10
1953:7	1954:5	1953:8	1954:8
1957:8	1958:4	1957:8	1958:4
1960:4	1961:2	1960:5	1961:2
1969:12	1970:11	1969:10	1970:11
1973:11	1975:3	1973:11	1975:7
1980:1	1980:7	1980:3	1980:7
1981:7	1982:11	1981:7	1983:4
1990:7		1990:7	

TABLE 3
Prewar Peaks and Troughs

<u>NBER Reference Dates</u>		<u>New Dates</u>	
<u>Peak</u>	<u>Trough</u>	<u>Peak</u>	<u>Trough</u>
1887:3	1888:4	1887:2	1887:7
1890:7	1891:5		
1893:1	1894:6	1893:1	1894:2
1895:12	1897:6	1895:12	1897:1
1899:6	1900:12	1900:4	1900:12
1902:9	1904:8	1903:7	1904:3
1907:5	1908:6	1907:7	1908:6
1910:1	1912:1	1910:1	1911:5
1913:1	1914:12	1914:6	1914:11
		1916:5	1917:1
1918:8	1919:3	1918:7	1919:3
1920:1	1921:7	1920:1	1921:7
1923:5	1924:7	1923:5	1924:7
1926:10	1927:11	1927:3	1927:12
1929:8	1933:3	1929:9	1932:7
1937:5	1938:6	1937:8	1938:6
		1939:12	1940:3

TABLE 4**Duration of Contractions and Expansions**

<u>Sample Period</u>	Average Length of Contractions (Months)		Average Length of Expansions (Months)	
	<u>NBER Dates</u>	<u>New Dates</u>	<u>NBER Dates</u>	<u>New Dates</u>
1887-1917	17.7	9.7	24.2	32.2
1918-1940	18.0	13.7	26.0	27.3
1948-1990	11.0	12.4	51.5	50.3

TABLE 5
Cycles Ranked According to Output Loss

<u>1887-1917</u>		<u>1918-1940</u>		<u>1948-1990</u>	
<u>Peak</u>	<u>Loss</u>	<u>Peak</u>	<u>Loss</u>	<u>Peak</u>	<u>Loss</u>
1916	0.46	1939	0.65	1980	0.44
1887	0.58	1927	0.68	1960	0.96
1914	0.75	1918	0.71	1969	1.05
1900	0.80	1923	1.89	1948	1.13
1903	1.16	1937	5.79	1953	1.25
1895	1.36	1920	6.64	1957	1.41
1910	1.53	1929	31.18	1981	1.67
1893	2.60			1973	2.39
1907	3.04				

Notes: The loss measure shows the cumulative loss in industrial production between the absolute peak associated with each cycle and the return to peak. To facilitate comparison, however, the dates of the peaks listed are the years corresponding to the final turning points identified by the algorithm and given in Tables 2 and 3.

FIGURE 1

Effect of Detrending on Turning Points

Logarithm of the Index
of Industrial Production

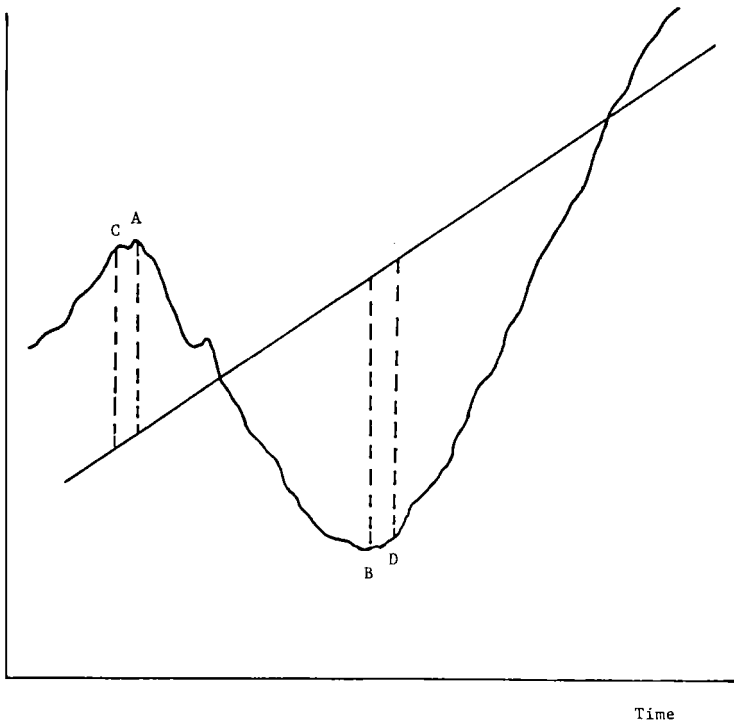


FIGURE 2

The Concept of Output Loss

Logarithm of the Index
of Industrial Production

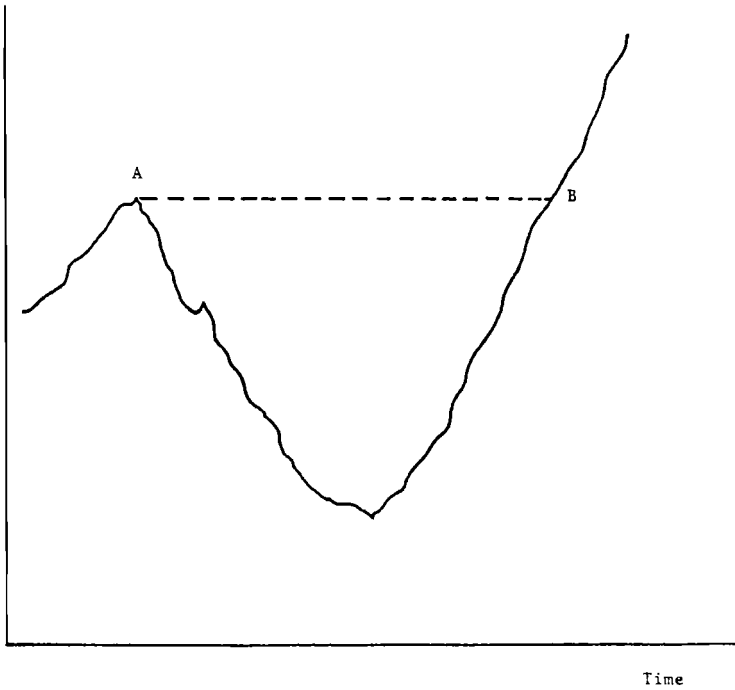
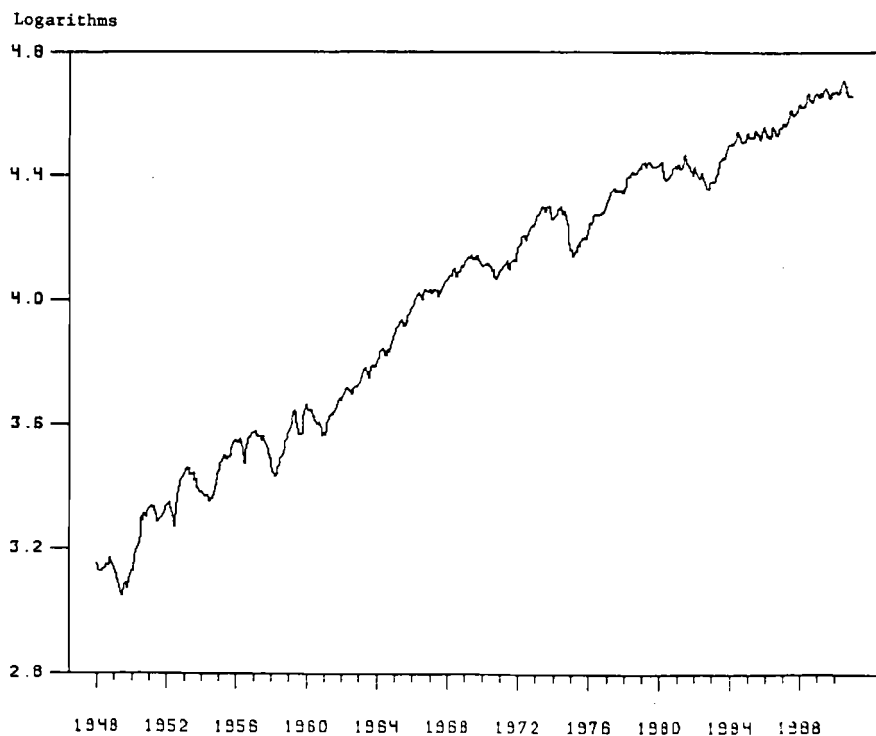


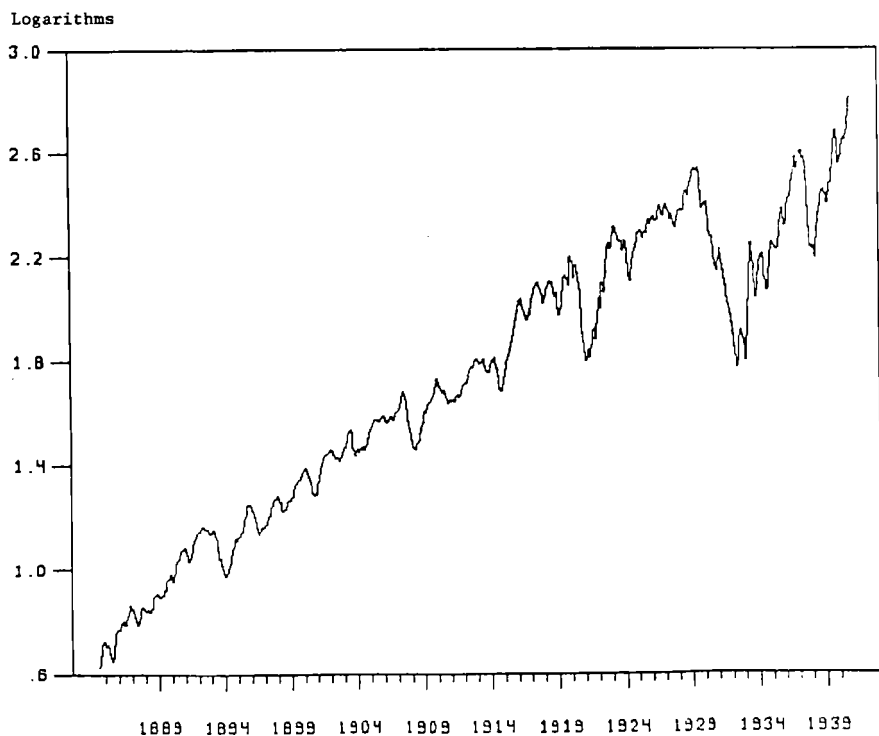
FIGURE 3
FRB Index of Industrial Production
(1948-1990)



Notes: This series is seasonally adjusted. See the text for a discussion of the source of the series and the seasonal adjustment.

FIGURE 4

**Index of Industrial Production
(1884-1940)**



Notes: For 1884-1918 this series is the seasonally adjusted, smoothed version of the Miron-Romer index described in the text. For 1919-1940 this series is the FRB index of industrial production which I have also seasonally adjusted.