Dating Turning Points in the Business Cycle

Michael D. Boldin

Journal of Business, Volume 67, Issue 1 (Jan., 1994), 97-131.

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

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

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

Journal of Business
©1994 University of Chicago Press

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

©2001 JSTOR

http://uk.jstor.org/
Tue Sep 11 12:14:53 2001

Michael D. Boldin
Federal Reserve Bank of New York

Dating Turning Points in the Business Cycle*

I review and evaluate procedures for dating peaks and troughs of the business cycle. Five different methods are judged on their historical performance, ease of replication, clarity, timeliness, flexibility, and the validity of underlying assumptions. In addition, I examine the 1990–92 period in detail where two newer dating methods are especially insightful because they use statistically rigorous methods to derive turning point probabilities, and, thereby, the inherent uncertainty in the data can be recognized. I also conclude that additional research is needed on this topic to resolve some important controversies about business-cycle characteristics.

I. Introduction

Business-cycle turning points, dates that the economy switches from expansion to recession and vice versa, have always been a source of interest in macroeconomic research. These turning points are often used to compare the amplitude and other relevant characteristics of different cycles, to identify unique periods, and to test competing economic theories. Most important, business-cycle dates continue to play a role in efforts to determine the causes of recessions and to design public policy that would prevent or at least limit the duration and impact of economic downturns.[1]

Numerous methods to identify turning points, with varying degrees of technical sophistication,

* The views expressed here are mine and do not necessarily reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System. I am grateful to many colleagues at the bank, especially Charles Steindel and Patricia Mosser, who provided useful comments and suggestions. George Green at the Bureau of Economic Analysis, Jim Stock, and the research staff at the National Bureau of Economic Research were also very helpful in providing data and clarifying some important issues. Cindy Silverio provided excellent assistance in this project. I alone am responsible for any errors or misinterpretations.

1. Consider the Gramm-Rudman-Hollings law, where Section 254, entitled "Special Procedures in the Event of a Recession," provides for suspension of mandated deficit ceilings when real economic growth is negative in 2 consecutive quarters. See Zarnowitz and Moore (1991) for a detailed discussion of this section of the law.

(*Journal of Business*, 1994, vol. 67, no. 1)
© 1994 by The University of Chicago. All rights reserved.
0021-9398/94/6701-0002$01.50

have been developed over the years. Like much of applied economics, the dating of turning points has been practiced as a craft as well as a science. This article reviews the most popular procedures, evaluates their historical performance, and discusses the validity of the underlying assumptions. Each method is also judged on ease of replication, clarity, timeliness, and flexibility.

Since labels of peak and trough do not necessarily affect fundamental conditions in the economy, it is sometimes argued that the dating of recessions and expansions is not germane to basic economic research. The best response to this concern is that business cycles are undeniably real-world phenomena. Different recessionary periods seem to have more in common with each other than they do with the expansions that surround them. Therefore, economists often use turning point dates to discern how the economy operates in these two separate regimes.

One problem is that official or consensus dates are not available when the study of the business-cycle data is most timely, and previous research has tended to focus on forecasting turning points. However, quick confirmation that a recession has begun or ended is almost as elusive as accurate predictions of future business-cycle trends. Therefore, this article argues for giving greater attention to real-time business-cycle dating methods.

To help judge the usefulness of the different dating methods, each was applied to the 1990–92 period, when the course of the economy was very uncertain. The National Bureau of Economic Research (NBER), considered by many to be the official source of the U.S. business-cycle chronology, declared a peak in July 1990 and a trough in March 1991. Some of the alternative dating methods suggest that the recession continued into mid-1992. Compared to past cycles, such wide disagreements are unprecedented, as trough dates are usually much clearer than peaks. The confusion can be traced to the fact that employment growth continued to be very weak after other measures of economic activity, such as industrial production, had clear troughs in 1991. This suggests that an important structural change may have occurred.

Besides providing unique insights into the current economic situation, the various dating procedures and their problems point out areas of further interesting research. Issues that pertain to both the statistical and theoretical validity of dating procedures are highlighted. They include the appropriateness of various detrending procedures, evidence of breaks and asymmetry in the data, and the implications of models that reject the popular view that both recessions and expansions usually die of old age (duration dependence). Continued research is definitely needed to improve our ability to understand both turning points and the more general business-cycle phenomenon.

II. Business-Cycle Terms and Concepts

Before examining and comparing different business-cycle dating methods, it is useful to review some important terms and concepts. Usually, economists define recessions and expansions in accordance with Burns and Mitchell (1946, p. 3): "Business cycles are a type of fluctuation found in aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle; this sequence is recurrent but not periodic; in duration business cycles vary from more than one year to ten to twelve years; they are not divisible into shorter cycles of similar character with amplitudes approximating their own." Stages of the cycle are inferred primarily from the level of economic activity. Turning points are called *peaks*—the period immediately preceding a decline in real activity, or *recession*—and *troughs*—the period immediately preceding an upturn, or *expansion*. In this article, distinctions based on growth-cycle concepts, a related theory that attempts to identify periods of above and below average growth in the economy, are not made. Therefore, the dating of recessions is sufficient since periods of rapid recovery are not separated from latter stages of an expansion.

One controversy with the Burns and Mitchell definition is whether it allows for a *single-index* view that assumes a common cyclical component in the fluctuations of many different data series. Some economists reject the single-index view and instead rely on *reference-cycle* dates that summarize general movements in economic activity.[2] There is also considerable debate on the use and abuse of detrending techniques that are used to determine peak and trough dates. Terms such as *nonstationarity, unit roots,* and *random walks* are used in the more technical research. The main controversy pertains to assumptions that the level of real output is *trend reverting* and follows a well-defined path. It is well-known that spurious relationships are less likely when growth rates are used, but important information can be lost when the effects of trend reversion are ignored.

Single-index models usually assume that business-cycle fluctuations are *extrinsic* phenomena and do not affect assumptions about the underlying structure of the economy. Other models and dating techniques

2. For example, Sargent and Sims (1977) refer to Koopmans's (1947) critique of Burns and Mitchell's work in justifying the single-index view as a useful first approximation. In contrast, Zarnowitz and Moore (1986) emphasize the cyclical conformity and coherence of numerous variables. This disagreement turns out to be less important than expected, however, since a "true" index is not observable and there is considerable noise in constructed proxies.

explore the importance of breaks in trends. For example, one of the dating methods estimates distinct changes in the equation that describes a particular time series. Therefore, this model has two statistical regimes, and cycles are *intrinsic*. An important feature of this intrinsic view is that business cycles can be *asymmetric,* such that recessions and expansions are not mirror images of each other. In contrast, constant parameter and single-index models invariably assume that positive and negative shocks have symmetric effects.

Another characteristic that is related to the extrinsic/intrinsic dichotomy, because it pertains to regime switches, is *duration dependence*. This view holds that the likelihood of a turning point increases the longer a particular regime goes on. The duration dependence view is seen in much of the economic analysis on Wall Street, where it is often claimed that a turning point is imminent just because the current recession or expansion has persisted for a longer than average time.

III. Turning Point Dating Methods

The business-cycle dating procedures that are discussed below are (*a*) the NBER Business Cycle Dating Committee approach, (*b*) gross domestic product (GDP) rules of thumb, (*c*) peaks and troughs of the Commerce Department's business cycle indicators, (*d*) Stock and Watson's experimental business cycle indicators, and (*e*) Markov switching models.[3] One standardization I make is to use only data that was available as of January 1993 for the main analysis. The effects of revisions before and after this date are noted and discussed when important.

To qualify as a useful dating technique, each procedure needs (1) careful and clear documentation of the data that is examined and (2) a means to distinguish recessions from expansions. The second step, known as a *pattern recognition problem,* does not have to be inflexible. In fact, flexibility can be useful since business cycles are irregular in periodicity and each recession has some unique factors and conditions. At the same time, dating procedures should be replicable and not seem arbitrary under both cursory and more careful expert examination.

The NBER approach is a good example of a flexible dating procedure. This private organization has a long history of research on U.S. business cycles, and their business-cycle chronology is considered by most economists to be the official source for peak and trough dates. However, it should not be assumed that any particular set of turning point dates are "correct," and one of the main purposes of this article

3. I also looked at two sequential dating procedures—Zarnowitz and Moore (1982) and Neftci (1982)—that were not very useful in confirming turning point dates.

is to investigate the merits of alternative dating methods. For example, simple rules based on n consecutive periods of upward/downward movement are commonly applied to either GDP or monthly indices of economic activity. However, there exists considerable confusion over whether these inflexible rules define "official" business-cycle turning points. Stock and Watson and Markov switching model methods take much more sophisticated approaches that have not been fully discussed or explored in the literature.

A. NBER Business Cycle Dating Committee

The current members of the NBER Business Cycle Dating Committee are William Branson, Martin Feldstein, Benjamin Friedman, Robert J. Gordon, Robert E. Hall (chair), Geoffrey Moore, and Victor Zarnowitz. There has been almost no turnover in membership since the committee's inception in 1980.[4] Prior to 1980, the NBER was also the main source for business-cycle dates with Moore and Zarnowitz being principals in work that expanded on the pioneering studies of Burns and Mitchell.

The committee's method for selecting turning point dates is pragmatic since it requires a consensus among members who tend to use different methods to analyze macroeconomic conditions and trends. Nonetheless, NBER reports, press accounts, and articles by the members of the committee provide a good understanding of the most important factors. (For example, see Zerwitz [1989].) Under the NBER approach, peak and trough dates are selected by looking for clear changes in both the trend and level of economic activity. Numerous data series that are believed to be coincidental with the aggregate economy are analyzed, and clusterings of turning points are used to set the reference cycle dates. The committee's decision process seems to closely adhere to Burns and Mitchell's concept of the business cycle, which (1) requires full cycles to last over 1 year and is skeptical of those lasting less than 2 years, and (2) chooses later (as opposed to earlier) turning point dates, both in periods of flatness and of multiple spikes (unless the spikes show a clear downward or upward pattern).[5] To provide consistency, comparisons are also made to patterns observed at previous NBER dates. Although the committee decisions can be criticized as slow in forthcoming, this approach avoids premature and false calls.

1990–92 period. In April 1991, the committee designated July 1990

4. The only change has been with Martin Feldstein's position. He was an original member but was replaced in 1981 (with Eli Shapiro) when he became head of President Reagan's Council of Economic Advisors. He returned to the dating committee in the late 1980s.

5. See Burns and Mitchell (1946, pp. 56–59), and note that the second criterion can result in the selection of a peak (trough) that does not necessarily correspond to the highest (lowest) point in an expansion (recession) phase.

as the peak, and in the official account of the decision, Hall (1991/92) explained that this month was a compromise. Figure 1 shows that as of January 1993 the four series Hall considers most important in dating decisions had individual peaks of April (income), March (employment), August (manufacturing and retail sales), and September (industrial production, or IP).[6]

In December 1992, the NBER committee announced that the recession's trough occurred in March 1991. Table 1 shows that the 20-month delay was long compared to their previous decisions. Before dating the 1980 and 1982 troughs, the committee took 12 and 8 months, respectively. The committee justified the delay as necessary in making sure that real GDP had clearly exceeded its 1990 high point and that there was a true expansion.[7] March 1991 was chosen because IP started a clear upward movement in April and there was a leveling off in employment at this time.[8] Figure 2, which provides comparisons to recessions in 1970, 1975, 1980, and 1982, shows that IP turned up sharply near each of the previous trough designations. In this respect, the March 1991 date seems inconsistent with the previous NBER decisions.[9]

Figure 2 shows one problem with this date, however. Within 12 months of previous troughs, all four series have clearly rebounded and are close to or exceed their prior peaks. Twenty months after March 1991, only sales and IP showed a true rebound and were close to 1990 levels. The sustained weakness in employment and income was unique. Therefore, placing greater weight on employment and income data, it can be argued that a general expansion did not begin in 1991. This interpretation of the data is also more consistent with Burns and Mitchell's criteria that tend to choose later turning point dates.

Despite economic sluggishness in the mid-1991 to mid-1992 period, none of the four series showed a true double-dip. The declines in IP and sales in late 1991 were modest and quickly reversed. Statements about a double-dip (made largely in the business press) seem based on

6. The original income peak was in July, and the employment peak was June. Later data revisions slightly altered these patterns. In fact, revisions after January 1993 moved the employment peak back to June and the trough was changed to February 1992. Also, the depth of the downturn was lessened in more recent data releases (relative to that shown in fig. 1).

7. It is also likely that political acumen kept the committee from making an announcement before the 1992 presidential election.

8. This explanation is based on earlier press statements by both Hall and Moore.

9. To evaluate the possibility that the low point in January 1992 is a more appropriate trough date, I applied the algorithm that Romer (1992) devised to replicate the post–World War II NBER selection process. Her procedure requires that a second distinct trough be chosen over an earlier, lower value unless the cumulative gain is greater than 11%. The cumulative gain between March 1991 and January 1992 was over 20%, and, therefore, the earlier date is chosen.

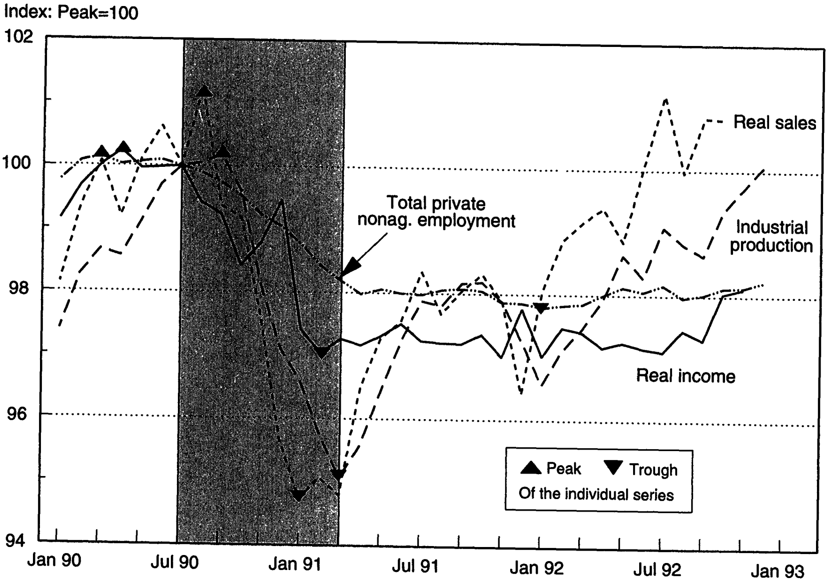

FIG. 1.—The recession of 1990–91

TABLE 1 **NBER Announcements of Business-Cycle Dates**

Date	Type	Announced	Months of: Delay	Duration
January 1980	peak	June 1980	6	6 recession
July 1980	trough	July 1981	12	12 expansion
July 1981	peak	January 1982	7	16 recession
November 1982	trough	July 1983	8	110 expansion
July 1990	peak	April 1981	9	8 recession
March 1991	trough	December 1992	20	. . .

comparisons of upturns in series like IP to later downturns in employment. Such an apples-versus-oranges comparison makes this view unappealing. Comparisons to the 1980–83 period, in which the committee identified two distinct recessions, also are useful in judging the plausibility of a double-dip decision. In deciding that a brief, but true, expansion began in mid-1980, the NBER committee argued that the recovery clearly followed normal patterns.[10] While the 1991–92 period was unusual in many respects, the identification of a second separate NBER-type recession cannot be made.

10. Source: *NBER Reporter* (1981/82, 1983). Hall (1991) also cautions against identifying false troughs such as the unsustained upturn that started in January 1982.

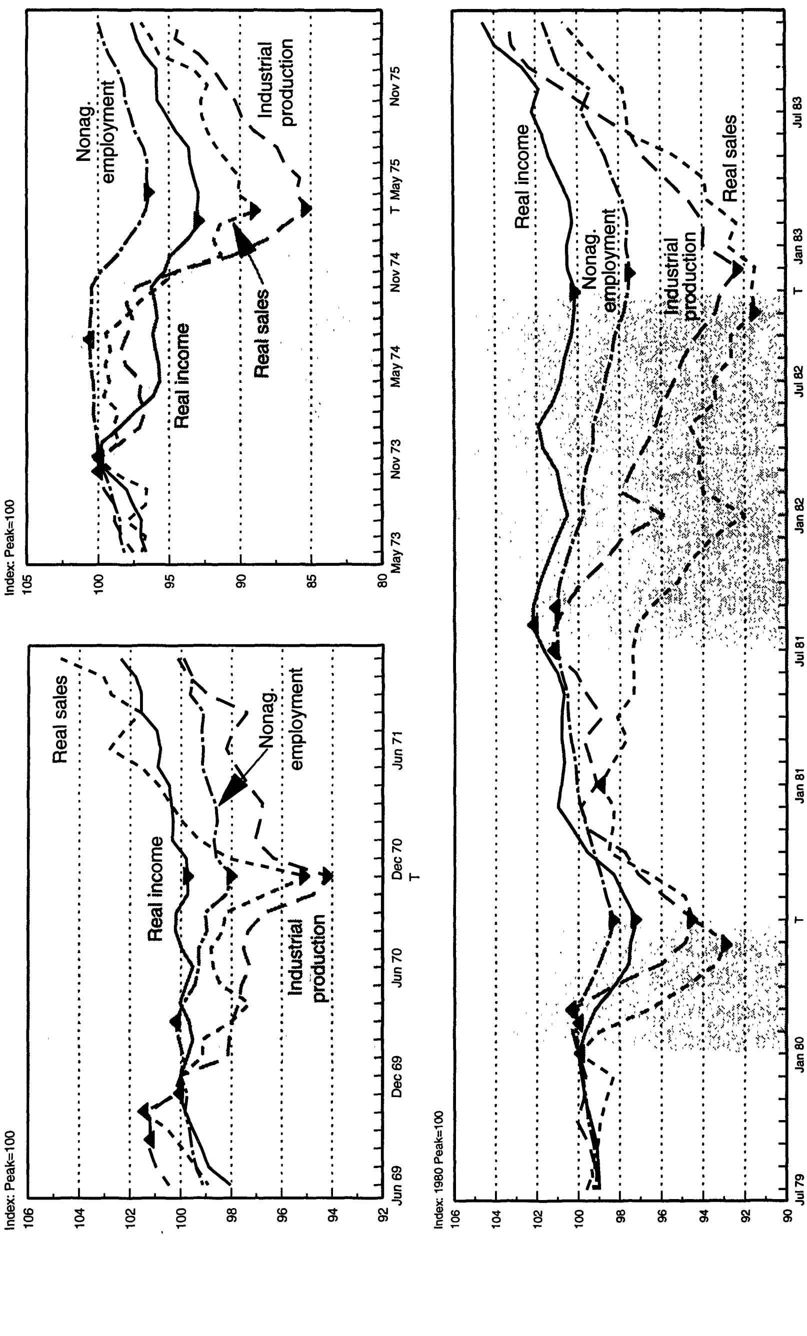

FIG. 2.—Prior recessions: 1969–70, 1973–75, 1980, 1981–82

B. GDP Growth Rules

Many economists believe that 2 consecutive quarters of negative output define the start of an official recession. Figure 3 shows that the correspondence between a "2-quarter GDP" rule and NBER recession dates is not very exact, however. One clear disagreement is the very short recession of 1980, when the percentage decrease in the second quarter (1980:2) was the largest in the postwar period and growth in 1980:3 was only 0.1%. Also, the barely positive growth rate (0.4%) in 1960:3 caused the 1960–61 recession to be missed. By strictly applying a 2-quarter rule, only three of the six NBER recession dates since 1960 are captured (using a tolerance of plus or minus 1 quarter).[11]

Figure 3 clearly demonstrates that peaks and troughs designations are harder to make with growth rates than they are with levels data (such as in figs. 1 and 2). Also, while zero growth as the cutoff point in defining the start of a recession is reasonable, the appropriateness of a symmetric criterion for the beginning of an expansion is less clear. McNees (1991) evaluated quarterly based output rules and argued that a 1% cutoff model is more useful. Figure 3 shows that an alternative asymmetric rule that sets peaks whenever there are 2 out of 3 quarters of negative growth and subsequent troughs whenever there are 2 out of 3 quarters of growth greater than 2.5% would more closely match NBER dates. The 2.5% cutoff for the beginning of an expansion is useful since growth should be higher than average after a recession ends. In fact, very high GDP growth (usually over 5%) confirms that the NBER trough dates mark distinct changes in the economy.

1990–92 period. The NBER's July 1990 peak date is consistent with the beginning of the latest recession being marked by 2 consecutive quarters of negative GDP growth. Positive growth started in 1991:2, which would support placing the trough in spring 1991 as the NBER committee did. While it is true that there have never been two consecutive increases in GDP before an NBER trough, it is just as noteworthy that weak growth in the last 3 quarters of 1991 (between 0.6% and 1.7%) contrasts strongly with the pattern of prior recoveries. The 2-out-of-3-quarters rule, which looks for growth over 2.5%, places the end of the recession in 1991:4.

C. Commerce Department's BEA Indicators

The Commerce Department's Bureau of Economic Analysis (BEA) computes three popular series for business-cycle analysis: the coinci-

11. The gross national product (GNP) has typically been used for this method, but I keep to the new emphasis on GDP. Nonetheless, the conclusions are quite robust to using GNP.

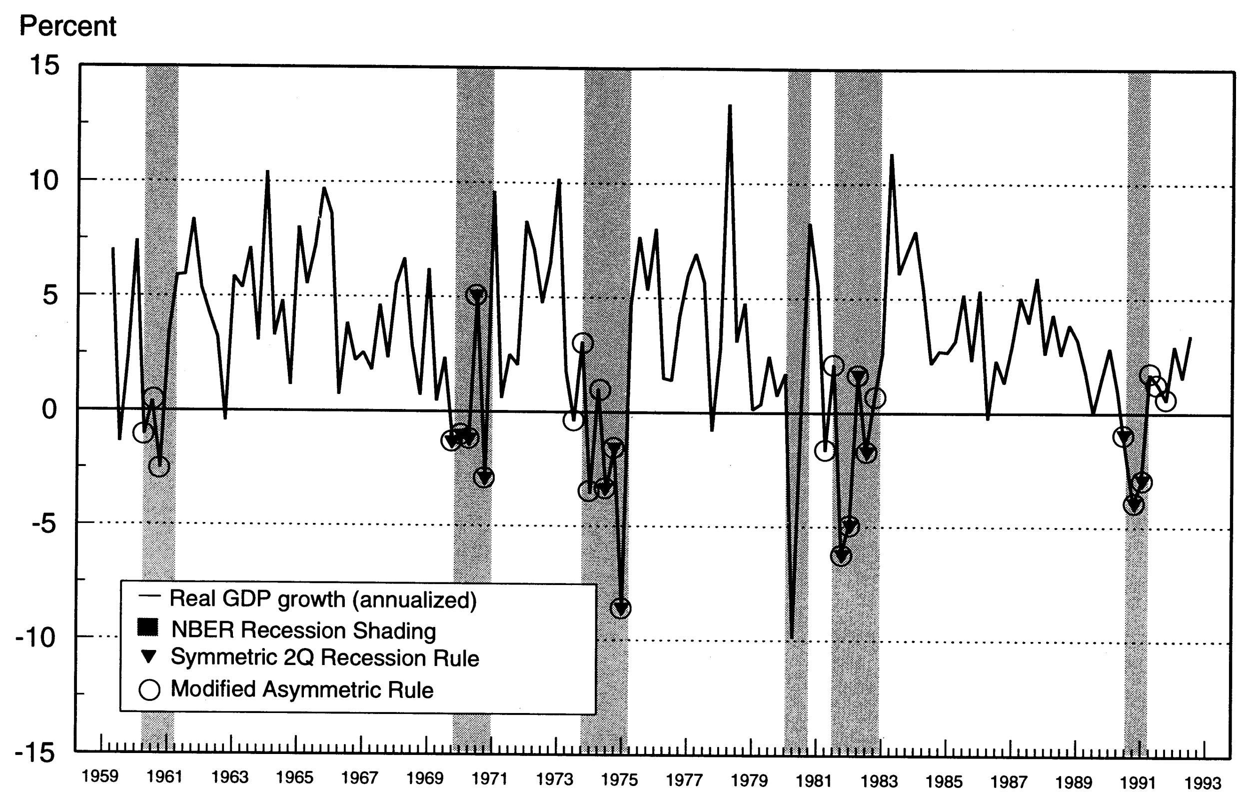

FIG. 3.—Real gross domestic product (GDP) growth and turning point rules. 2-quarter rule: recessions begin in the first quarter that GDP growth turns negative and the following quarter's growth rate is also negative. Symmetrically, subsequent expansions begin with GDP growth turning positive 2 quarters in a row. Modified rule: recessions begin in the first quarter that GDP growth turns negative if growth is also negative in at least 1 of the next 2 quarters. Subsequent expansions begin only when GDP growth (annualized) exceeds 2.5% and in at least one of the next two quarters, growth again exceeds 2.5%.

dent (CI), leading (LI), and lagging (LgI) indicators. The construction methodology for these series is based on research that began with Burns and Mitchell and was continued by other NBER economists. The monthly percentage change in each indicator is a weighted average of growth in individual components:

$$g_t = \sum_{i}^{n} \beta_i g_{i,t} + a.$$

By summarizing various data sources into one series, these indicators follow the single-index view. The weights (β_i) are inversely related to the estimated volatility of each component's growth rate ($g_{i,t}$). Adjustments (a) are also used to give each indicator the same underlying trend as GDP, which is currently assumed to be 3.0%.[12]

The CI series uses the same four series highlighted by Hall (with only a slightly different definition for employment). Not surprisingly, figure 4 shows that peaks and troughs occur almost exactly at NBER dates. It is noteworthy that employment data receive the greatest weight, almost three times as much as the IP series, which explains the relative smoothness of the CI. It is also evident in figure 4 that LI and LgI, which have different components, are much more volatile but show appropriate lead and lag relationships.

The BEA has not established a clear rule for choosing the peaks and troughs they have designated for each series. Indeed, there seems to be considerable flexibility as some dates are not even (local) minimums or maximum values in the surrounding 12-month period. This is not to say that the selections are arbitrary.[13] I applied a three consecutive increase/decrease rule to CI, and table 2 shows reasonable matches to the official BEA turning points as well as the NBER dates. However, a modified rule is even better. Table 2 shows results from requiring 3-out-of-4 months in a new direction and a 2% cumulative change before declaring a turning point. However, the last column in the table shows significantly greater delays in confirming a turning point.

I also applied the same turning point rules to the LI, LgI, and coincident-to-lagging and leading-to-coincident ratios. The results are not shown because obviously false signals were common, and consistent lead or lag times to peaks and troughs in general economic activity were not evident. This is the general finding with these types of indicators—the signals of a turning point are very noisy and unreliable.

12. See Hertzberg and Beckman (1989).

13. The BEA clearly attempts to be consistent with and build on the Burns and Mitchell tradition of business-cycle analysis. Also, they acknowledge the contributions of Moore and Zarnowitz to the latest revision of their methods (as described in Hertzberg and Beckman [1989]). Bry and Boschan (1971) showed that a computer program can match reasonable judgmental selections of peaks and troughs like these with a high degree of accuracy.

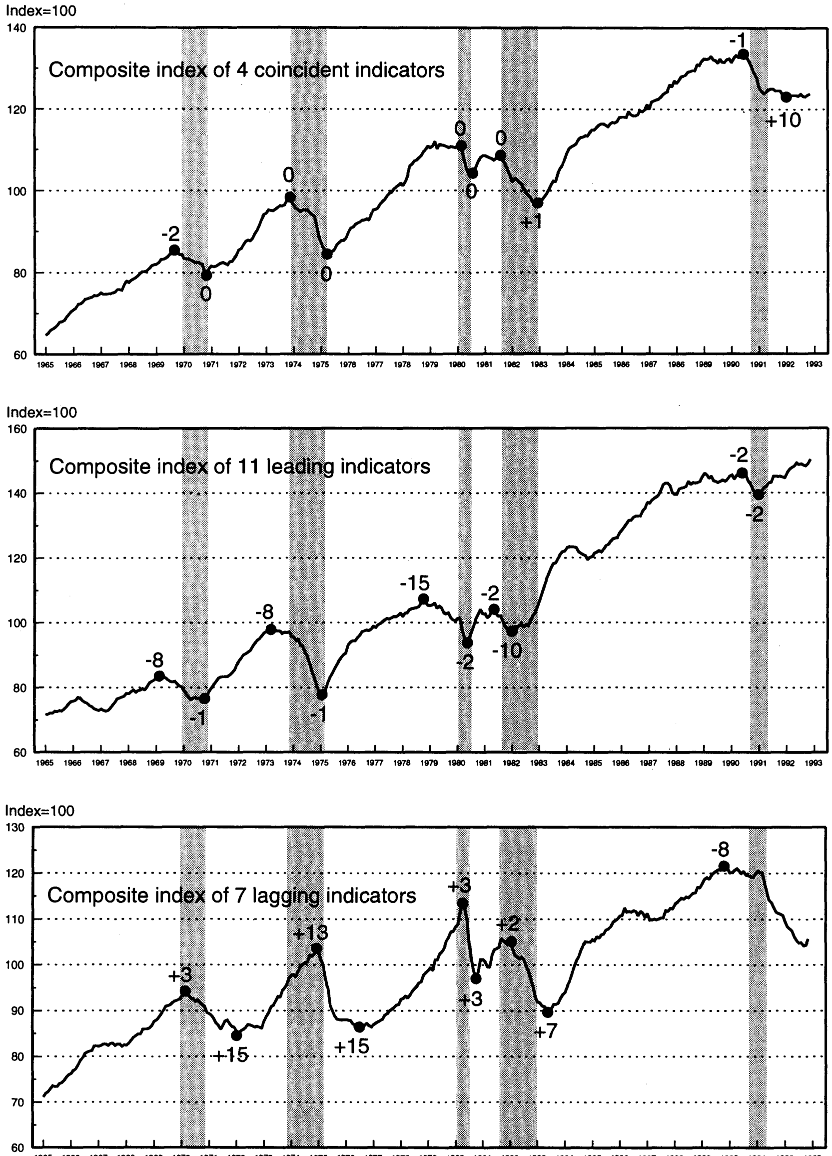

FIG. 4.—Bureau of Economic Analysis business cycle indices. Dots indicate official peak and trough destinations with number of months leading (−) or lagging (+) National Bureau of Economic Research dates.

TABLE 2 The Bureau of Economic Analysis's Coincident Indicator

BEA Official Turning Points		Different from NBER	Three-Consecutive-Changes Rule				Modified Three-out-of-Four Rule			
				Date Difference				Date Difference		
Type	Date		Date	BEA	NBER	% Change	Date	BEA	NBER	Delay
Peak	October 1969	−2	March 1970	5	3	−1.08	October 1969	0	−2	4
Trough	November 1970	0	February 1971	3	3	1.11	November 1970	0	0	4
Peak	November 1973	0	November 1973	0	0	−2.75	November 1973	0	0	3
Trough	March 1975	0	May 1975	2	2	2.24	March 1975	0	0	5
Peak	January 1980	0	May 1979	−8	−8	−.54	March 1979	−10	−10	12
Trough	July 1980	0	July 1980	0	0	2.88	July 1980	0	0	3
Peak	July 1981	0	January 1981	−6	−6	−.65	July 1981	0	0	4
Trough	December 1982	1	February 1983	2	3	2.66	December 1982	0	1	4
Peak	June 1990	−1	April 1989	−14	−15	−1.20	June 1990	0	−1	4
Trough	January 1992	+10	March 1991	−10	0	.89	no signal			
			October 1991	False peak?		−1.44				

SOURCE.—BEA data release of January 1993 (series ends in November 1992); official 1990–92 peak and trough designations were made in March 1993.
NOTE.—BEA = Bureau of Economic Analysis; NBER = National Bureau of Economic Research.

1990–92 period. In the initial BEA releases, a possible CI peak was first evident in mid-1990. However, table 2 shows that later revisions to the CI, driven by revisions to the underlying data, caused the three consecutive decreases rule to place a peak back in April 1989. The modified 3-out-of-4 rule rejects this date because the subsequent decrease was less than 2% and instead chooses June 1990. The three consecutive increases rule then finds a trough in March 1991 with a subsequent peak only 7 months later. For data that were available as of January 1993, yielding a CI series up to November 1992, there were no further signals from the two rules. In early 1993, the BEA designated June 1990 as the CI peak and February 1992 as the CI trough.

The 1991–92 CI patterns do not really qualify as a true double-dip since it is based on flatness in the series rather than a clear change in direction. The CI level in March 1991 (123.9), short-term peak in July 1991 (125.0), subsequent low point in January 1992 (122.8), and November 1992 value (123.7) are all within 2% of each other. Therefore, the modified rule did not find a trough. Also, the LI series, which peaked very early in 1989, had a clear trough in January 1991, but 2 years later it was only 8% above this low point. This pattern contrasts strongly with previous periods that saw increases of over 20% in the LI during the first 18 months of a recovery.

The flatness in the CI from mid-1991 through 1992 does not mean the economy was not growing. Since the trend adjustment for CI is −2.2% per year (without the correction it tends to grow 2.2% faster than GDP), it can decline even when growth in the underlying components is positive. This property has led some economists to doubt the CI's usefulness, and the BEA is studying alternative methods for constructing this indicator.[14] Still, prior recoveries had higher than average growth rates, and the negative trend adjustment was not seen as a source of bias. The adjusted CI shows that the economy did not follow a typical recovery pattern.

One problem in drawing conclusions about the real-time usefulness of the CI is that the patterns in figure 4 and results from table 2 are based on revised data. Figure 5 shows that first releases of the CI give very volatile signals about the economy.[15] From the beginning of 1990

14. See Green and Beckman (1992). The problem with the current methodology is that the component weights (β_i) are not constrained to sum to one, but instead equal 1.83. This overweighs cyclical movements, relative to trend factors. An alternative method, which the BEA introduced in late 1993, corrects this problem and results in an apparent trough in March 1991.

15. A substantial downward revision was made in August when revisions in the components back to 1987 were incorporated into the calculations of CI. Usually revisions are made only to the past 6 months. This revision affected the level of the CI much more than it did the cyclical pattern. To keep the graphs of the series more comparable, I adjusted all of the earlier releases down 1.2 points to match the level of the postrevision CI in January 1990.

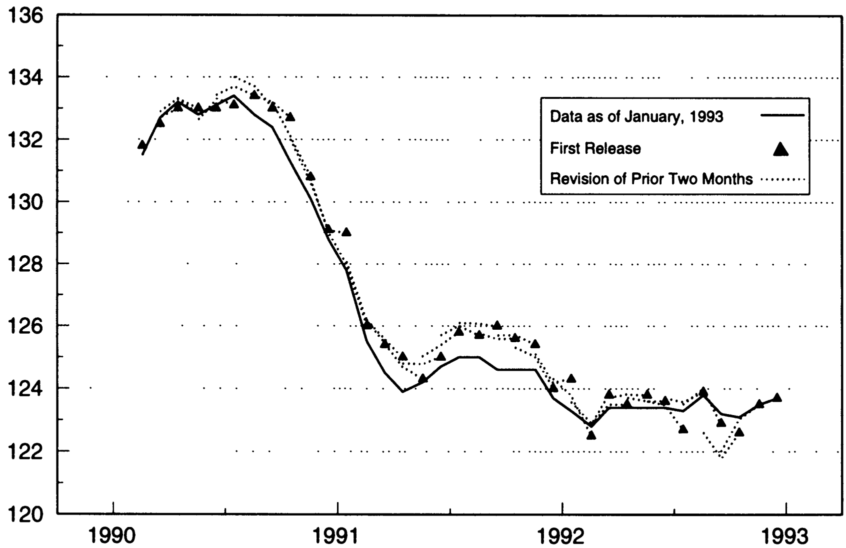

FIG. 5.—Effects of revisions on the Bureau of Economic Analysis coincident indicator.

to the end of 1992, revisions led to four reversals (over 10% of the time) of an initially reported increase or decrease. These patterns are consistent with research by Diebold and Rudebusch (1991) that shows similar problems with the LI and should make economists cautious about drawing trends from preliminary data. However, the variance around levels is reasonably limited, which is consistent with Zarnowitz's (1982, 1992) research on the revision problem.

D. Stock and Watson's Experimental Business-Cycle Indices

In an NBER supported project, Stock and Watson (S-W; 1989, 1991) developed an experimental coincident indicator (XCI) of the business cycle. Following the single-index view, a common cyclical element was assumed to fully describe underlying business-cycle trends. Although this factor is unobservable, by assuming it affects several macroeconomic series in the same manner,

$$x_{i,t} = b_0 + b_1 c_t + u_t,$$

a Kalman filtering algorithm can be used to derive the expected value of c_t based on actual observations of the x_i series.

Stock and Watson used roughly the same data sources as the BEA's CI. The only real exception is that hours worked is substituted for the employment count. They were also careful to transform each series into growth rates because of concern that the levels were not stationary. The derived common cyclical growth factor was then transformed back into levels to yield the series XCI that is shown in figure 6. This chart shows that while movements in the CI and XCI are often very similar, the latter had a clear upturn in early 1991. This difference can be traced to the fact that the S-W index is more sensitive to changes in the IP than it is to employment, while the opposite is true for the CI. Also, since mid-1980, the XCI series more closely matches movements in aggregate output, which is shown in figure 6 using an index that was created by interpolating GDP to a monthly frequency.[16]

Another interesting feature of S-W's work is their construction of an experimental leading indicator (XLI). This index forecasts growth in the XCI 6 months into the future using various leading indicator series in a simple regression format. This well-defined relationship between the XCI and XLI can be interpreted as a more formal and statistically rigorous representation of previous attempts to find a stable predictive relationship between the BEA's LI and output (measured using the CI, IP, unemployment rate, or GNP).[17]

Instead of using simple recession and expansion rules, Stock and Watson convert growth rates of XCI into another indicator that measures the probability that a recession is occurring. Similar exercises are performed with the XLI to yield recession forecasts. This part of the S-W procedure is very complicated and computationally demanding since it entails defining the characteristics of a recession and calculating the odds that a certain pattern is consistent with this definition. Stock and Watson used a stochastic recession rule that requires 6 months of low or negative c_t values. This pattern recognition procedure is Stock and Watson's interpretation of the NBER approach, and, since it does not include information about the level of economic activity, its appropriateness and efficiency is open to debate.[18]

1990–92 period. Every month since 1989 the NBER releases an update of the XCI and XLI calculations, providing a true out-of-sample test of the S-W procedure. The XCI first turned down sufficiently to

16. As reported above, the poorer performance of the CI has prompted the BEA to reconsider its methodology. The BEA credits Stock and Watson with pointing out problems with the trend adjustments. No explicit trend adjustments are made to the XCI, which has a lower mean monthly growth rate than GDP, but higher volatility gives it a similar compound growth rate. (See Stock and Watson's [1992] discussion on p. 13 and table 2.2.)

17. For example, Zarnowitz and Moore (1981) used smoothed growth rates of the CI and LI to yield sequential signals (both predictions and confirmations) of a turning point but relied on more arbitrary definitions of the threshold growth rates for a new phase in the business cycle. I applied this method to the 1984–92 period and found relatively poor performance. Also see Auerbach (1982) for a finding that has been consistently repeated: while the LI seems quite useful in predicting output fluctuations, in-sample forecasts can be improved by changing the component weights but the out-of-sample record of an "optimized" leading indicator deteriorates considerably.

18. The S-W calculations build on the work of Wecker (1979) and Kling (1987), but little guidance is available on the selection of the exact criteria.

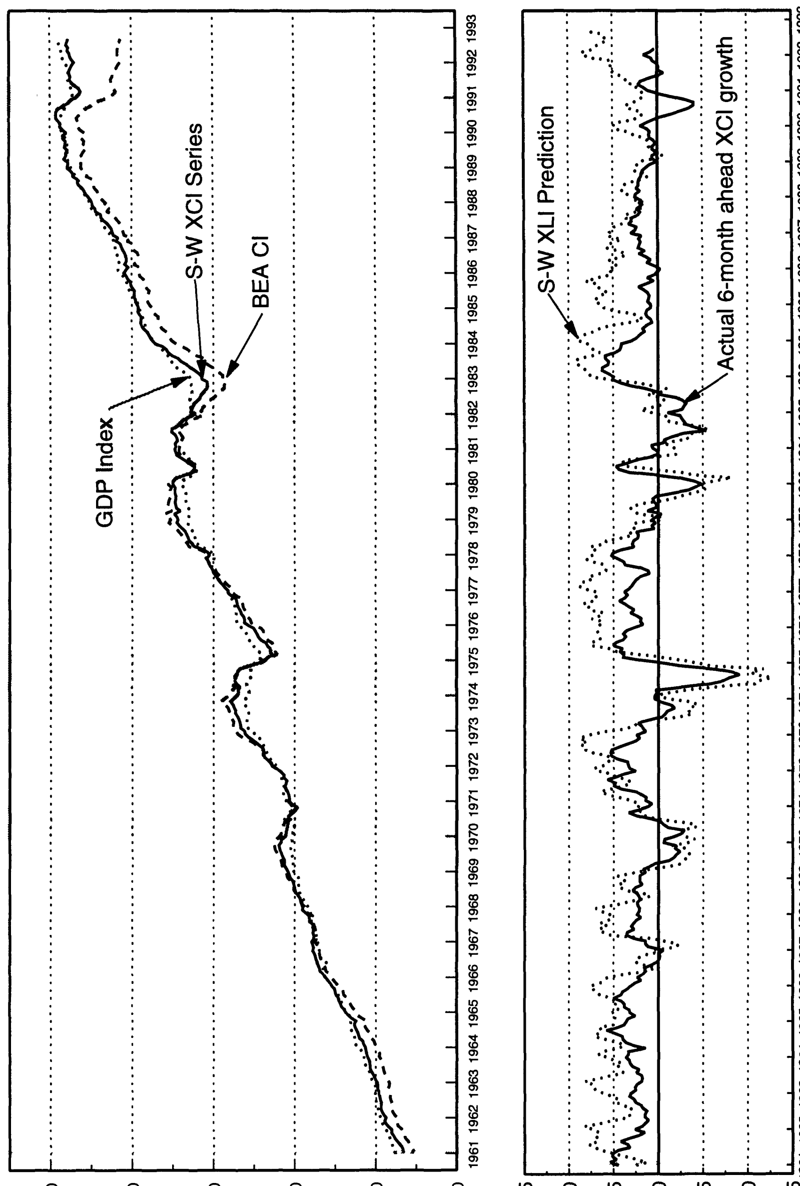

FIG. 6.—Stock-Watson experimental business-cycle indices

define a recession (under the S-W criteria) after October 1991. Unfortunately, the XLI completely missed this downturn.[19] (See the bottom panel of fig. 6, where expected growth was over 3% throughout 1990 and 1991.) In subsequent analysis, Stock and Watson (1992) traced this problem to the heavy influence of financial variables, most notably the commercial paper/T-bill spread, that were useful in predicting previous recessions but were neutral in the latest period. Noting this problem, Stock and Watson then constructed a second leading index, XLI-2, that uses only nonfinancial variables. The XLI-2 gave a recession forecast in July 1990 and a relatively weak signal of an upturn in early 1991. In May 1991, the XCI moved out of recession, but during the next year and a half, the XLI consistently overpredicted the XCI's growth. (Stock and Watson do not publish XLI-2 values, only the recession probabilities from this index that have ranged between 17% and 36% since mid-1991.) In any event, the XCI recession probabilities rose above the 50% criteria in November 1990 and fell below this line in May 1991.[20] This record is reasonably consistent with the NBER decisions they were designed to capture.

E. *Markov Switching Model for Unemployment*

The last dating method I review has only been recently developed and explored. This framework is based on Markov switching models (MSMs) that were introduced to business-cycle analysis by Hamilton (1989). Hamilton's MSM specification captured distinct periods of high (positive) and low (negative) growth in quarterly GNP, and Boldin (1992) expanded this research to monthly data. In these types of models, low-growth periods can be labeled *recessions* and high-growth periods are *expansions,* formalizing the belief that business cycles are intrinsic to the economy.

With the MSM framework, the data make all the dating decisions, or, more correctly, the probability that a particular time period was in recession is calculated. Explicit references to traditional business-cycle dates are not needed. Instead, it is possible to simultaneously derive the characteristics and dates for recessions and expansions. This feature contrasts with the methods discussed above that decide the differences between expansions and recessions before deciding the turning points. (For more details, see the Appendix; or Boldin [1992].)

19. Interestingly, Sims (1989) and Zarnowitz and Braun (1989) in commenting on Stock and Watson (1989) worried about an overreliance on interest rates and spreads and underrepresentation of real commitments by businesses. The latter pair noted that the LI "has a more comprehensive coverage based on a longer historical experience (and) this could prove an advantage over time inasmuch as the causes of business cycles may vary" (Zarnowitz and Braun 1989, p. 406).

20. The source is an NBER distributed newsletter, "Stock and Watson Indicator Report," that provides monthly updates of the S-W indices.

I estimated an MSM for monthly unemployment rates, over the 1960–89 period, using equations of the form

$$un_t = \alpha_{0,s(t)} + \alpha_{1,s(t)} un_{t-1} + \alpha_{2,s(t)} un_{t-2} + e_t, \quad s(t) = 1, 2.$$

It is assumed that there are two distinct regimes or equations—one for recessions and one for expansions. Only one operates at each time period and therefore the sequence $s(1), s(2), \ldots, s(T)$ defines the business-cycle chronology. Table 3 shows the estimated parameters of the model. The recession equation predicts that unemployment will continually rise while the expansion equation predicts slow convergence to an equilibrium rate of around 4% (computed by solving for $un_t = un_{t-1} = un_{t-2}$).

Two transition probabilities are also estimated to provide the odds of switching from one regime to the other:

$$q_{12} = \text{prob}(s(t) = 2 | s(t-1) = 1)$$

and

$$q_{21} = \text{prob}(s(t) = 1 | s(t-1) = 2).$$

TABLE 3 **Markov Switching Model for the Unemployment Rate (1960–89 Sample Period)**

$$un_t = \alpha_{0,s(t)} + \alpha_{1,s(t)} un_{t-1} + \alpha_{2,s(t)} un_{t-2} + e_t, \quad e_t \sim N(0, \sigma_{s(t)})$$

$$q_{i,j} = \text{prob}(s(t) = j | s(t-1) = i)$$

$$p_{0i} = \text{prob}(s(0) = i)$$

	Regimes (s)	
	Recession (1)	Expansion (2)
$\alpha_{0,s}$	.1126	.1172
	(.1162)	(.0340)
$\alpha_{1,s}$	1.0129	.7265
	(.1230)	(.0603)
$\alpha_{2,s}$	−.0026	.2436
	(.1270)	(.0590)
σ_s	.2154	.1375
	(.0197)	(.0070)
$q_{1,s}$	.9137	.0863
	(.0408)	. . .
$q_{2,s}$	.0248	.9752
	(.0118)	. . .
p_{0s}	.2230	.7770
Log likelihood	138.4771	
No. of observations	360	

NOTE.—These are maximum likelihood estimates with standard errors in parentheses.

With this model, the estimated probability of a recession ending is 8.6% (for any month of a recession). As with all MSMs, the switches are independent of all other factors, including how long a particular regime has lasted. The corresponding probability of switching from expansion to recession is much lower, 2.5%. These switching probabilities are consistent with the view that a long expansion is more likely than a long recession. Markov switching models do not assume duration dependence, however, and a new recession or expansion could continue indefinitely.

A useful option with switching models is to sequentially reevaluate the probabilities as new data become available. Not only can the probability of a recession be calculated for the most recent month, revisions to the probabilities of prior months can be made from the longer sample. Because business cycles are persistent, these revisions recognize that data observed after a particular month can provide valuable information about the likelihood of a particular regime on that date.

Figure 7 shows the unemployment rate data and the MSM's recession probabilities. While the probabilities acknowledge that data can never tell us definite or infallible recession and expansion dates, it is possible to select the "most-likely" sequence of regimes and turning point dates. These are also in figure 7, and the results closely match NBER dates, which would surprise economists who believe unemployment rates lag the general cycle. In fact, the upper panel of the chart shows that unemployment rates tend to rise before NBER peaks and fall after NBER troughs. Because two separate equations are estimated, MSMs can flexibly capture this asymmetry at turning points.

1990–92 period. Figure 7 also shows recession probabilities for the 1990–92 period, which are out-of-sample results. Here, the model clearly picks up the economic downturn that began in mid-1990. Table 4 provides additional information about this period in the form of sequential reevaluations of the monthly recession probabilities as new data become available. The fourth column is most noteworthy since it provides recession probabilities that are concurrent with the last month of the data release. Here, recession probabilities first rose above 50% after release of the August 1990 unemployment rate. The month that directly succeeds the most-likely peak is July 1990, and its recession probability increased to 80% after the release of data for early 1991.

The MSM recession probabilities were generally over 85% from December 1990 to June 1991. For the full-sample probabilities, a spike down to 66% occurred in July 1991. Preliminary recession probabilities of less than 50% were first seen after the August 1991 release, and a low of 16% was given when the November data were released. These declines can be traced to the fact that unemployment was no longer on a clear upward path. Subsequent increases revised this premature signal of an expansion, and the full-sample probability that November

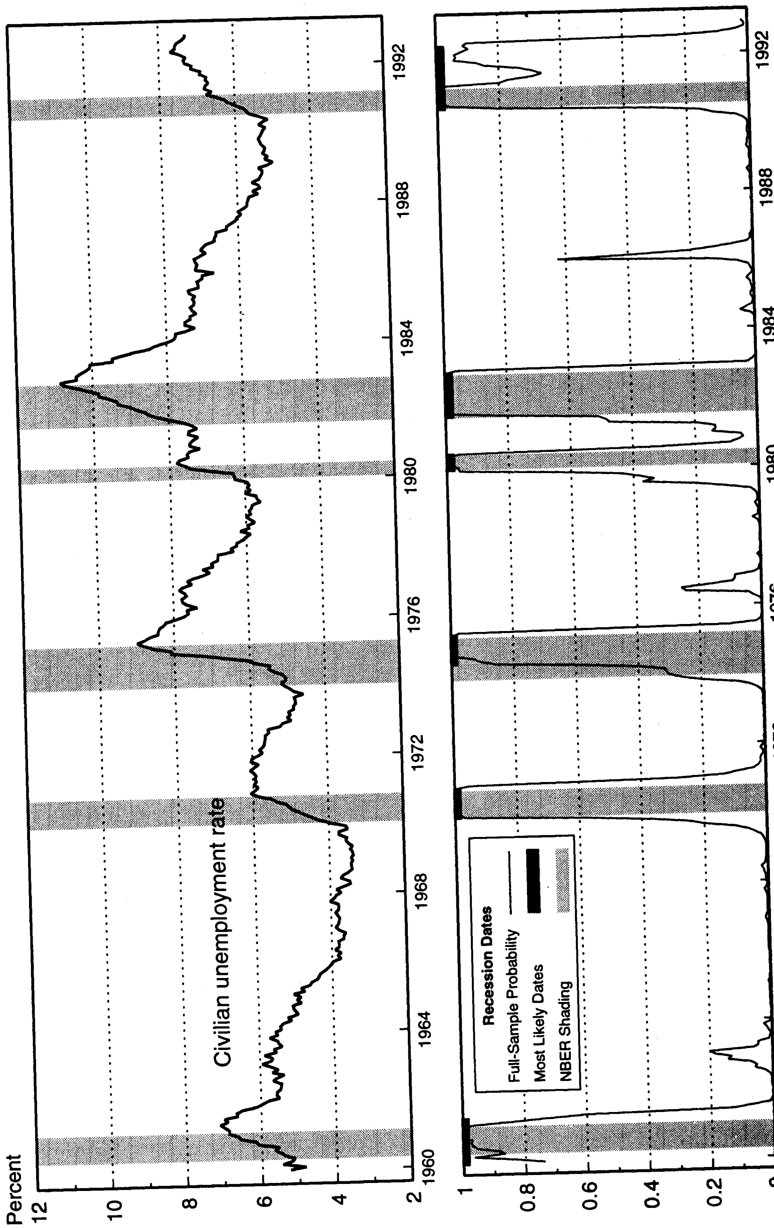

FIG. 7.—Markov switching model for unemployment

TABLE 4 **Markov Switching Model Recession Probabilities with Sequential Revisions**

Month	Unemployment Rate		Recession Probabilities $p(s_t = 1 \mid y_{t+k}, \ldots, y_1)$ $k =$							Revised for Full Sample
	Original	Revised	0	1	2	3	4	5	6	
January 1990	5.28	5.30	.0025	.0013	.0007	.0009	.0006	.0005	.0006	.0097
February 1990	5.27	5.27	.0125	.0051	.0074	.0039	.0028	.0041	.0093	.0192
March 1990	5.18	5.20	.0125	.0200	.0087	.0054	.0094	.0258	.0279	.0436
April 1990	5.36	5.38	.0593	.0237	.0134	.0260	.0774	.0841	.0830	.1175
May 1990	5.27	5.27	.0262	.0113	.0294	.1037	.1132	.1117	.1463	.1550
June 1990	5.19	5.14	.0180	.0546	.2043	.2235	.2204	.2903	.3125	.3151
July 1990	5.43	5.38	.1317	.5043	.5521	.5445	.7183	.7736	.7842	.8020
August 1990	5.64	5.62	.5879	.6452	.6361	.8442	.9103	.9230	.9445	.9660
September 1990	5.72	5.70	.6132	.6028	.8422	.9183	.9329	.9576	.9590	.9772
October 1990	5.78	5.81	.5577	.8361	.9245	.9415	.9702	.9718	.9716	.9864
November 1990	5.99	6.00	.8373	.9380	.9573	.9900	.9919	.9916	.9918	.9962
December 1990	6.13	6.17	.9168	.9451	.9931	.9957	.9953	.9957	.9958	.9979
January 1991	6.22	6.26	.8959	.9911	.9964	.9956	.9964	.9964	.9962	.9969
February 1991	6.50	6.50	.9898	.9992	.9978	.9991	.9992	.9988	.9986	.9984
March 1991	6.72	6.76	.9949	.9816	.9940	.9955	.9909	.9897	.9872	.9929
April 1991	6.57	6.61	.6640	.8929	.9196	.8356	.8134	.7666	.7845	.7972
May 1991	6.81	6.80	.8882	.9201	.8199	.7934	.7375	.7588	.7322	.7716
June 1991	6.86	6.77	.8690	.6973	.6519	.5561	.5926	.5470	.6863	.6920

July 1991	6.77	6.70	.5230	.4505	.2975	.3559	.2830	.5054	.4845	.6628
August 1991	6.82	6.77	.4080	.2356	.3013	.2192	.4698	.4463	.6610	.6946
September 1991	6.77	6.75	.1969	.2695	.1788	.4558	.4298	.6671	.6493	.7370
October 1991	6.88	6.92	.2824	.1792	.4943	.4646	.7345	.7143	.6241	.8453
November 1991	6.86	6.90	.1614	.5107	.4778	.7771	.7547	.6746	.7955	.8759
December 1991	7.08	7.15	.5745	.5363	.8843	.8582	.7651	.9057	.9382	.9452
January 1992	7.08	7.13	.4917	.8901	.8602	.7537	.9146	.9518	.9515	.9381
February 1992	7.32	7.31	.9075	.8741	.7549	.9349	.9766	.9763	.9761	.9440
March 1992	7.30	7.34	.7688	.5450	.8830	.9612	.9606	.9602	.9599	.9186
April 1992	7.22	7.28	.4154	.8575	.9598	.9591	.9585	.9581	.9579	.8979
May 1992	7.47	7.45	.8830	.9962	.9953	.9947	.9943	.9941	.9941	.9153
June 1992	7.82	7.69	.9996	.9982	.9972	.9964	.9961	.9960	.9024	.9024
July 1992	7.65	7.56	.6292	.4203	.2580	.1941	.1823	.2339		.2339
August 1992	7.61	7.55	.3393	.1493	.0744	.0606	.0969			.0969
September 1992	7.52	7.50	.1167	.0341	.0189	.0349				.0349
October 1992	7.35	7.38	.0276	.0089	.0133					.0133
November 1992	7.23	7.30	.0126	.0114						.0114
December 1992		7.27	.0196							.0196

NOTE.—At each date the recession probabilities were calculated for that date and also updated for all past dates, based on the new information. The probabilities in columns labeled by a k value are for that date based on data up to k periods ahead. For example, in the July 1990 row, the .1317 value in the $k = 0$ column is the probability that July was a recession month based on data only up until that month. The next column, $k = 1$, shows an increase in the recession probability to .5043 (for July 1990), when data from August 1990 is added. The last column provides the recession probabilities from the full sample (all of the data in col. 3) that are also used in fig. 7. Also note that the third column shows the effect of revisions to the official seasonal adjustments that was made with the December 1992 release.

1991 was a recession month moved to over 85% as additional months of data became available.

The MSM results imply that the recession continued into 1992. Declines in the unemployment rate in July and August then moved the probabilities to less than 50%. With the originally reported November 1992 rate falling to 7.23%, the July recession probability (which became the most-likely first month of the recovery) fell to 18%. With the December 1992 data, seasonal adjustment factors were changed back to 1988, but the effect was marginal in deriving a recession probability of less than 2% for the end of 1992.[21] In any event, the MSM results confirm that the 1990–92 period was quite unusual and that large revisions in recession probabilities are possible as trends in the data become more clear.

IV. Comparison and Evaluation of the Methods

The dating procedures discussed above can be evaluated by various criteria. Table 5 provides my rankings in six categories: historical performance, ease of replication, clarity, timeliness, flexibility, and validity. Others may come up with a different ordering, especially with respect to validity that I define by the reasonableness of the underlying theoretical and statistical assumptions. Below, I discuss in detail the reasons for these rankings. In summary, none of the procedures are clearly superior. In many areas, high marks are given to the S-W and MSM procedures because they use the most systematic and internally consistent means to evaluate business-cycle data. The NBER approach is also given above average marks because of its flexibility and historical status.

Historical performance. Commonly, the historical performance of a dating method is checked by how closely it matches NBER dates. But this implies that the NBER dates are indisputable, and even members of the NBER committee would admit that this is too strong an interpretation. While a complete performance ranking that avoids selecting the true set of turning point dates cannot be made, we can note that the different dating methods usually find turning points that are within a few months of each other. Therefore, all render usable sets of peak and trough dates. However, some of the methods are less reliable because they tend to vary randomly between early and late designations (compared to consensus views).

While the choice of individual NBER turning point dates can be debated, the general concept has withstood years of controversy, dramatic changes in the economy, and obviously premature conclusions that the business cycle was dead. Acceptance of these dates seems

21. Seasonal adjustments were also revised in 1991, but the effects were immaterial.

TABLE 5 **Ranking of Dating Methods**

Criteria	NBER Committee	GDP Rules of Thumb	CI Rules	S-W Procedure	MSM for Unemployment
Performance	1	5	4	2	3
Replication	3	1	2	5	4
Clarity	1	3	2	5	4
Timeliness	5	4	3	2	1
Flexibility	1	5	4	3	2
Validity	3	5	4	2	1
Average	2.3	3.8	3.2	3.2	2.5
No. of criteria ≤ 2	3	1	2	3	3

NOTE.—NBER = National Bureau of Economic Research; GDP = gross domestic product; CI = coincident indicators; S-W = Stock and Watson; MSM = Markov switching model. The orders are 1-best to 5-worst. Replication of the NBER approach is based only on replication of the methods used by members of the committee. The average values should not be taken too seriously since they are based on equal weights. The count of rankings in the top two provides a better gauge of better than average scores.

based not on inertia but on time-tested respect for the continuity and objectivity of the NBER. Therefore, the NBER methods get the highest marks for historical performance. In contrast, simple peak and trough rules for GDP and CI are less reliable and consistent than is commonly thought. Adjustments that add flexibility to these rules yield more dependable turning point decisions.

If we ignore 1990–92, which seems reasonable considering the uniqueness of the period, the S-W and MSM records are about equal. While these newer dating methods have not been subjected to comprehensive historical tests, at least on an ex post basis, both capture the general chronology of the NBER dates. Since they are particularly efficient in processing data and summarizing the important information, future work is warranted with these models. It seems particularly worthwhile to make comparisons where both methods use the same set of data. Directions that this course of research may take are discussed below under the topic of theoretical validity.

Ease of replication. Simple rules of thumb (e.g., *n* consecutive periods of change in a new direction) for GDP and BEA indicators are simple to carry out and have the greatest ease of replication. For the NBER approach, exact replication is not possible, but most business economists have experience conducting the types of analysis that are in published accounts of the dating decisions. The two most technically demanding methods are the MSM and S-W procedures. The latter was a very time-consuming project that has not yet been duplicated by others or greatly modified by Stock and Watson.[22] Markov switching

22. Stock and Watson have made their computer programs available to others, and the BEA has experimented with the procedure. However, I know of no full replications and only one significant alteration (Crone 1994).

models are also computationally intensive as they rely on nonlinear estimation techniques that are less well known than conventional regression procedures. Still, the multiple regime aspects have strong intuitive appeal for many areas of economic research, and the use of MSMs is growing.

Clarity. Clarity, or how easily the output of a statistical procedure can be interpreted, is another important criterion. In this respect, some economists prefer simple dating rules because they provide the most clear and discrete signals of a turning point. However, this advantage may be a red herring that overstates the information content of the data. In other words, a clear signal may grossly overstate the odds that a turning point date will hold up under further analysis. By using probability measures, the S-W and MSM procedures handle this inherent uncertainty more realistically. Despite the fact that they do not provide absolute and discrete separations of recession and expansion periods, the results are still useful.

Timeliness. Simplicity and the ease of replication certainly influence another important consideration—the timeliness of a dating procedure. However, well-designed computer programs can greatly narrow any handicaps from updating and adding new data to the computationally intensive procedures. Therefore, only the NBER decisions have a significant lag.

In considering timeliness, it is clear that business cycle fluctuations are evident in monthly series that provide finer, and therefore more useful, detail than GDP data. Consensus on this point is seen by the fact that only one method relies on quarterly data. The only additional data concerns are based on delays in availability and revision effects. Therefore, the MSM for unemployment is most timely since unemployment data is released only a few days after the month's end. Revisions to these data are usually quite small and only occur at the annual updates to the seasonal adjustment factors. Releases of the BEA indicators are typically over a month later and are subject to relatively large revisions.[23] Stock and Watson release an update of their indices about the same time but do not discuss revision effects. While none of the dating procedures really addressed the revision problem, I was surprised to find that these changes had little consequence on the ex post evaluations of turning point decisions.

Flexibility. A remaining controversy is whether flexibility should be allowed in deciding turning points. Renshaw (1991), in arguing for fixed rules, complains that "the NBER should either explain its method of identifying economic recessions, so they can be dated by

23. Usually, personal income and sales data are available at the end of the next month, and official calculations of the BEA series are released within a week. It is possible to replicate the calculations and thereby save a few days.

newspaper reporters and other subscribers to the Survey of Current Business, or forgo the responsibility of dating business peaks and troughs altogether." He fears that the current NBER process is too arbitrary and "constitutes mysticism, not science" (p. 58).

Although NBER decisions involve some judgment, they are clearly not arbitrary. Some controversy is unavoidable when flexibility is allowed, and debates about individual periods do not diminish the usefulness of the NBER chronology. In fact, disputable judgments are incorporated in the supposedly fixed rules. For example, for the 2-quarter GDP rule, a decision was made to use 2 instead of 3 quarters and to focus on GDP and not individual components, such as consumption and investment. It should be recognized that preliminary decisions are always necessary in designing and implementing statistical procedures, and the efficient processing of information will always include some discretion.

It is not clear how fixed rules capture the complicated dynamics of a recession. Why use the word "recession" when it is also correct to say "the period after 2 consecutive quarterly declines in GDP"? The obvious answer is that recessions represent a broader phenomenon. But this argues for basing business-cycle dates on a broad source of information and for not applying an infexible rule to one data series. Also, modifications of the most popular rules (to incorporate flexibility) seemed to yield better turning point decisions. These points motivate the NBER committee's approach that could be modeled as a form of "fuzzy" logic that has a true scientific basis. In this regard, we would expect seven other, similarly educated economists (given the same directions) to come up with dates that would be similarly accepted as capturing the business-cycle phenomenon.[24]

Validity. Both the S-W and MSM dating procedures follow a middle course between fixed decision rules and subjective judgment. Here, recession probabilities summarize the amount of confidence that can be placed on a specific turning point date. In addition, both procedures are consistent with modern economic research that stresses the importance of coherent and plausible empirical models. Therefore, they receive the highest marks for theoretical and statistical validity. Both procedures show that a well-thought-out modeling structure can be very useful in finding turning points.

Even though the S-W and MSM procedures have rigorous statistical foundations, they are not based on predetermined or undebatable

24. A spirited and historically interesting debate between Cloos (1963*a*, 1963*b*) and Zarnowitz (1963*a*, 1963*b*) shows that NBER methods and dates have never been accepted without controversy. Zarnowitz's conclusion that "one may wish to regard the NBER method as a 'proxy', but if so, it is a proxy for an ideal (a single precise measure of aggregate economic activity) which is not attainable" (1963*b*, p. 461) reasonably sums up the views of most economists.

structures. Other data series and specifications could be explored. In considering extensions, it is obvious that turning points are much more evident in the levels data. However, it is noteworthy that the S-W procedure, which only relies on growth-rate data, was able to closely match most NBER decisions. Still, it should be realized that even in cases where an econometric model should be estimated in growth-rate form, the turning point decisions do not have to ignore information about levels.[25]

In developing or modifying a dating procedure, besides choosing between levels and growth rate data, there are numerous other important decisions. For example, empirical modeling decisions should recognize that many views (both academic theories and popular, but less rigorously developed, concepts) of the business cycle are consistent with breaks in the data or regime switches. The most common econometric techniques (e.g., vector autoregressions and the concepts underlying S-W's procedure) follow the extrinsic view and ignore the possibility that recessions are unique episodes. This concern, and general dissatisfaction with models that arbitrarily separate trends from cycles, motivated Hamilton to propose the use of MSMs for empirical business-cycle research. He argued that trends and cycles are fundamentally intertwined and that business-cycle patterns result from changes in the trend. With a minimal number of assumptions, MSMs can capture the different statistical regimes that result from the intrinsic aspects of economic fluctuations.

The intrinsic/extrinsic distinction is also linked with assumptions about linearity and symmetry. Linear models, which remain the primary means of econometric analysis, imply symmetry—recessions and expansions are mirror images of each other and are only differentiated by the sign of random shocks. Recent empirical work has revived the once well-accepted view in the business-cycle literature that the irregular patterns in U.S. economic activity implied highly nonlinear dynamics and asymmetries.[26]

The main point to make with the asymmetry and nonlinear evidence is that ignoring these characteristics will tend to make dating procedures less efficient. It should be recognized that peaks and troughs can develop differently. Though all the dating methods could be modified to accommodate this view, this is rarely done. The exception

25. The random walk or unit root problem with levels data pertains to asymptotic conditions and does not rule out valid inferences about finite subsamples.

26. One caveat to this claim about linear models is that observed asymmetries can result from asymmetric shocks. Still, conventional estimation procedures assume spherical or symmetric disturbances that do not seem to be consistent with actual macroeconomic data. See Brunner (1992), Boldin (1992), and the *Journal of Applied Econometrics* (1992) for new research on business cycle asymmetries.

is with MSMs that explicitly model asymmetries by using different equations for recession and expansion. (It could also be argued that the NBER committee process implicitly allows for asymmetric decisions and that the NBER tradition follows the nonlinear/intrinsic view.)

Another interesting characteristic of MSMs is that turning points are modeled as random events with constant probabilities over the tenure of a regime. The constant switching probability assumption contrasts with many business-cycle theories that predict increasing likelihoods of a turning point or duration dependence. Evidence like the 110-month duration of the 1983–90 expansion and more systematic studies by Sichel (1989) and Diebold and Rudebusch (1990) support claims that the duration of a particular expansion does not help predict its end. The evidence against duration dependence in recessions is weaker, pointing out another possible asymmetric feature. One advantage of a Markov switching assumption for both recessions and expansions is that changes in trends are stressed, not the length of the particular regime, in making turning point decisions. Still, further work to incorporate realistic duration dependence features is warranted, but it should be recognized that the small number of full cycles (only eight in the 1952–92 period, according to the NBER chronology) will make estimates of duration dependence very uncertain.

In conclusion, future research on turning points needs to consider and resolve many statistical and theoretical concerns. Expanding on the S-W unobservable-component model and multiple-regime concepts of MSMs seems to hold the most promise. For instance, asymmetries and nonlinearities could be added to the equations used in the S-W procedure. Alternatively, an unobservable component effect could be incorporated in the MSM framework. This latter type of model could formally test the significance of intrinsic and extrinsic components of business cycles.

V. Review of 1990–92 Dating Results

Table 6 summarizes the peak and trough dates of the various dating methods for the 1990–92 period. Absolute peaks and troughs dates for the four series that Hall reported as most important to the NBER committee decisions are also denoted. Collectively, the various methods are consistent with the official NBER peak. However, there is wide variance in the trough date.

A key reason for uncertainty about the trough date is that, despite an upturn in production and sales, there was little change in private-sector hiring in 1991 and 1992. In fact, employment at the end of 1992 was only marginally higher than in early 1991. Therefore, it is not surprising

TABLE 6 1990–92 Peak and Trough Dates

Method	Peak	Trough
NBER committee	July 1990	March 1991
Industrial production*	September 1990	March 1991
Real income*	April 1990	February 1991
Nonagricultural employment*	March 1990	January 1992
Manufacturing and retail sales*	August 1990	January 1991
GDP 2-quarter rule	1990:1	1991:1
GDP asymmetric rule	1990:1	1991:4
CI official	June 1990	January 1992
CI 3-month rule†	June 1990	March 1991
CI 3/4-modified rule	June 1990	. . .
Stock-Watson	October 1990	April 1991
MSM for unemployment	June 1990	June 1992

NOTE.—All data are as of January 15, 1993. NBER = National Bureau of Economic Research; GDP = gross domestic product; CI = coincident indicator; MSM = Markov switching model; 1990:1 indicates the first quarter of 1990.

* For these series, the peak is the 1990 high point and the trough is the low point between July 1990 and November 1992 (i.e., no double-dip assumed).

† This rule also gives another peak in September 1991, and the October 1992 value is less than the March 1991 trough.

that it is difficult to identify a clear trough date. In addition, employment seemed to be a leading series in the downturn while production was lagging, which is a reversal of conventional patterns.

Many reasons have been given for the unusual patterns in the employment data, including complaints about seasonal factors, other adjustments, and revisions to the data. I examined these issues in some detail and concluded that the complaints are the result, not the source, of flatness in the 1991–92 period. With this flatness, any reasonable adjustment or revision could greatly affect the absolute trough date. In prior recovery periods, strong growth in employment overwhelmed these effects.

The most likely reason for the divergence of employment from patterns observed in prior recoveries is the restructuring and downsizing of many large corporations. The empirical problem with separating these short-term trends from conventional cyclical factors is that changes of this magnitude have not been observed before. Therefore, we cannot make comparisons to historical patterns. A theoretical obstacle is also daunting. It is likely that many firms accelerated their restructuring plans as aggregate demand fell and less manpower was needed. Therefore, the downsizing effects (i.e., secular trend) and cyclical factors are not independent. Efforts to disentangle these two effects would depend greatly on assumptions about the peak and trough dates.

There are some additional features about the recent period that the dating procedures point out. First, since early 1991 there have been increases in output and productivity without corresponding growth in employment and wages. This is evidence against real business-cycle theories that predict a stronger productivity-to-wage-to-employment link. However, rising inflation and countercyclical monetary restraint do not seem to have preceded the downturn, which conflicts with conventional monetary explanations of business fluctuations. Also, efforts by the Federal Reserve to kick start the economy have been less potent than many expected. In this regard, problems with credit supply and tighter lending standards by banks have been reported that warrant further exploration. In summary, the dating procedures help identify new and unusual trends even in periods when a consensus was not reached. Also, the uncertainty that is inherent in choosing peaks and trough dates makes simplistic explanations of the recession less plausible.

VI. Conclusion

Noting the problems in dating turning points, especially in the recent period, it is clear that pragmatism should not be abandoned. It is unlikely that a be-all and end-all technique for dating business cycles can ever be developed. Different methods can play complementary roles in building a consensus. In this respect, the S-W and MSM results are most enlightening since they provide an explicit and coherent statistical framework to derive turning point probabilities. Summarizing the data in this way recognizes that uncertainty about turning points will always exist. Still, the importance of judgment in determining the data to consider and in setting up parsimonious or tractable models can never be avoided. The effects of revisions in the underlying data sources are also very difficult to incorporate in the dating procedures. Finally, further structural changes are likely to require modification to any statistical model of the economy.

With respect to the 1990–92 period, there was no true consensus on the trough date, which is surprising since usually these types of turning points are sharper than corresponding peaks. Some series and dating methods point to a trough in Spring 1991. Others, especially those based on employment data, suggest an unusually long recession that extended into 1992. Further analysis shows that a "double-dip" is not the culprit, and revisions to data can account for only a small fraction of the confusion. Efforts to date the peak and trough show that the recent period was quite unusual.

Appendix

Markov Switching Model Specification and Estimation

A very general MSM framework can be based on linear equations:

$$y_t = X_t\beta_{s(t)} + e_t, \quad e_t \sim N(0, \sigma_{s(t)}), s(t) = 1, 2, \ldots, n. \tag{A1}$$

Here, n distinct sets of independent variable coefficients and standard errors are hypothesized for each possible regime. The Markov switching assumption also requires the estimation of a first-order transition matrix, which is represented in the two-regime case by

$$Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix},$$

where $q_{ij} = \text{prob}(s(t+1) = j|s(t) = i)$ characterizes the switching probabilities. The likelihood function for this model, which measures how well a set of parameter estimates fits the data, is

$$L(y, X, \beta, \sigma, Q) = \sum_{s(T)=1}^{n} \cdots \sum_{s(1)=1}^{n} \sum_{s(0)=1}^{n} \left[\prod_{t=1}^{T} f_{s,t} q_{s(t-1),s(t)}\right] p_{0,s}.$$

Here the n equations of (A1) define probability density functions (pdf's) $f_{s,t} = f(y_t - X_t\beta_s, \sigma_s)$ for each regime at each time period, and normality is usually assumed. I also set the initial probabilities at the unconditional values that satisfy $p_0 = p_0Q$.

Efficient computation of this likelihood function is discussed in Boldin (1992). Since this is a nonlinear problem, a numerical optimization routine must be used to compute maximum likelihood estimators (MLEs). Still, the MLEs satisfy very intuitive formulas:

$$\beta_i = (X_i^{*\prime}X_i^*)^{-1}(X_i^{*\prime}y_i^*),$$

$$\sigma_i^2 = (e_i^{*\prime}e_i^*) \Big/ \left(\sum_t p_{i,t}^*\right),$$

and

$$q_{ij} = \left(\sum_t p_{ij,t}^*\right) \Big/ \left(\sum_t p_{i,t}^*\right).$$

The elements in the starred matrices above are $X_i^* = \{X_t \sqrt{p_{i,t}^*}\}$, $y_i^* = \{y_t \sqrt{p_{i,t}^*}\}$, and $e_i^* = \{(y_t - X_t\beta_i \sqrt{p_{i,t}^*}\}$, where

$$p_{i,t}^* = \text{prob}(s(t) = i|y) = L(s(t) = i, y, X; \beta_s, \sigma_s, Q)/L(y, X; \beta_s, \sigma_s, Q)$$

denotes the full-sample regime probability that is computed for each period using Bayes's Rule. (In the calculation of q_{ij}, more complicated but analogous full-sample probability measures, $p_{ij,t}^* = \text{prob}(s(t-1) = i, s(t) = j|y)$, are needed in the numerator.) These formulas show that the MLE solutions for β_i and σ_i are equivalent to estimates from weighted regressions that are based on the probability that an observation belongs to regime i. Here, the most important point is that the MLEs are completely consistent with the full-sample probabilities that in turn are calculated using the estimated parameters. This is the sense that data splitting can be considered endogenously based and provides a contrast with Neftci's (1982, 1984) work that employed similar switching concepts but did not derive optimal parameters for the recession and expansion equations.

In figure 7, the Viterbi algorithm was used to derive the most-likely sequence, and it is important to note that this sequence is not equivalent to making selections based on an individual period's probabilities being greater than 50%. Instead of making each period's regime selection in isolation (ignoring all the other selections), the Viterbi algorithm uses a joint selection process where all selections are consistent with each other. The difference can be seen by realizing that the most-likely sequence would maximize

$$\text{prob}(s(T), s(T-1), \ldots, s(1) \mid y, X, \beta, \sigma, Q),$$

and not $\Sigma_t\ \text{prob}(s(t) \mid y, X, \beta, \sigma, Q)$.

Boldin (1990) shows the better performance of an MSM for unemployment over Hamilton's (1989) specification for GNP growth. Following the work reported in Boldin (1992), the number of lagged dependent variables in the monthly based unemployment model specification was chosen by maximizing the Schwartz criteria that subtracts $\log(T = 360)/2$ times the number of estimated parameters from the likelihood function. I searched over 1–12 lags of unemployment as right-hand-side variables and considered (but rejected) a time trend. I computed unemployment rates directly from the reported level of seasonally adjusted unemployment and labor force because the official unemployment rate series are rounded to one decimal place and, therefore, could not yield normally distributed errors for the MSM equations.

Noting the difference with the latest NBER trough dates, I experimented using the full 1960–92 sample period and alternative MSM specifications for unemployment. Including a time trend (to recognize that average unemployment rates increased in the late 1970s and 1980s) sometimes resulted in a trough date in early 1991. However, the recession probabilities were still generally in the 20%–30% range throughout the end of 1991, and these results were not stable to small changes in the sample period. A much better fitting model was found by using a fuel price index and labor participation rates as explanatory variables in the recession and expansion equations. Again, the most-likely trough date was June 1992, and results from the simple MSM reported above are representative of the more robust alternative models.

I also applied the MSM framework to IP growth rates and found less clear business-cycle patterns. First, a two-regime model does not seem to be sufficient. Even when three or four regimes are assumed, high volatility in IP growth makes it difficult to differentiate each period's regime candidates. Future research that could solve this problem by using a four-regime, bivariate MSM for the levels of both IP and unemployment rates is planned.

References

Auerbach, A. 1982. The index of leading indicators: "Measurement without theory," thirty-five years later. *Review of Economics and Statistics* 64 (November): 589–95.

Boldin, M. 1990. Characterizing business cycles with a Markov switching model: Evidence of multiple equilibria. Research Paper no. 9037. New York: Federal Reserve Bank of New York.

Boldin, M. 1992. Using switching models to study business cycle asymmetries: I. Overview of methodology and application. Research Paper no. 9211. New York: Federal Reserve Bank of New York.

Brunner, A. 1992. Conditional asymmetries in real GNP: A semi-nonparametric approach. *Journal of Business and Economic Statistics* 10 (January): 65–72.

Bry, G., and Boschan, C. 1971. *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. New York: National Bureau of Economic Research.

Burns, A., and Mitchell, W. C. 1946. *Measuring Business Cycles*. New York: National Bureau of Economic Research.

Cloos, G. 1963*a*. How good are the National Bureau's reference dates? *Journal of Business* 36 (January): 14–32.

Cloos, G. 1963*b*. More on reference dates and leading indicators. *Journal of Business* 36 (July): 352–64.

Crone, T. 1994, in press. New indexes track the state of the states. *Business Review*.

Diebold, F., and Rudebusch, G. 1990. A nonparametric investigation of duration dependence in the American business cycle. *Journal of Political Economy* 98 (June): 596–616.

Diebold, F., and Rudebusch, G. 1991. Turning point prediction with the composite leading index: An ex ante analysis. In K. Lahiri and G. Moore (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records*. Cambridge: Cambridge University Press.

Green, G., and Beckman, B. 1992. The composite index of coincidental indicators and alternative coincident indexes. *Survey of Current Business* 72, no. 6 (June): 42–45.

Hall, R. 1991. Economic fluctuations. *NBER Reporter* (Summer), pp. 1–2.

Hall, R. 1991/92. The business cycle dating process. *NBER Reporter* (Winter), pp. 1–3.

Hamilton, J. 1989. A new approach to economic analysis of nonstationary time series. *Econometrica* 57 (March): 357–84.

Hertzberg, M., and Beckman, B. 1989. Business cycle indicators: Revised composite indexes. *Survey of Current Business* 72, no. 6 (January): 23–28.

Journal of Applied Econometrics. 1992. Special issue: Nonlinear dynamics and econometrics. Vol. 7, suppl. (December).

Kling, J. 1987. Predicting the turning points of business and economic time series. *Journal of Business* 60 (April): 201–38.

Koopmans, T. 1947. Measurement without theory. *Review of Economic Statistics* 29 (August): 161–72.

McNees, S. 1991. Forecasting cyclical turning points: The record of the past three recessions. In K. Lahiri and G. Moore (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records*. Cambridge: Cambridge University Press.

NBER Reporter. 1981–82. (Winter), pp. 23–24.

NBER Reporter. 1983. (Summer), p. 2.

Neftci, S. 1982. Optimal prediction of cyclical downturns. *Journal of Economic Dynamics and Control* 4 (August): 225–41.

Neftci, S. 1984. Are economic times series asymmetric over the business cycle? *Journal of Political Economy* 92 (April): 307–28.

Renshaw, E. 1991. Measuring recessions. *Challenge* 34 (March–April): 58–59.

Romer, C. 1992. Remeasuring business cycles. Working Paper no. 4150. Cambridge, Mass.: National Bureau of Economic Research.

Sargent, T., and Sims, C. 1977. Business cycle modelling without pretending to have too much a priori economic theory. In C. Sims (ed.), *New Methods in Business Cycle Research: Proceedings from a Conference*. Minneapolis: Federal Reserve Bank of Minneapolis.

Sichel, D. 1989. Business cycle duration dependence: A parametric approach. Economic Activity Working Paper no. 98. Washington, D.C.: Board of Governors of the Federal Reserve System.

Sims, C. 1989. Comment [on Stock and Watson 1989]. In O. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual—1989*, pp. 394–97. Cambridge, Mass.: MIT Press.

Stock, J., and Watson, M. 1989. New indexes of leading and coincidental economic indicators. In O. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual—1989,* pp. 351–94. Cambridge, Mass.: MIT Press.

Stock, J., and Watson, M. 1991. A probability model of coincident economic indicators. In K. Lahiri and G. Moore (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records*. Cambridge: Cambridge University Press.

Stock, J., and Watson, M. 1992. A procedure for predicting recessions with leading indicators: Econometric issues and recent experience. Working Paper no. 4014. Cambridge, Mass.: National Bureau of Economic Research.

Wecker, W. 1979. Predicting the turning points of a series. *Journal of Business* 52 (January): 35–50.

Zarnowitz, V. 1963*a*. On the dating of business cycles. *Journal of Business* 36 (April): 179–99.

Zarnowitz, V. 1963*b*. Cloos on reference dates and leading indicators: A comment. *Journal of Business* 36 (October): 461–63.

Zarnowitz, V. 1982. On the function, quality, and timeliness of economic information. *Journal of Business* 55 (January): 87–120.

Zarnowitz, V. 1992. *Business Cycles: Theory, History, Indicators, and Forecasting*. Chicago: University of Chicago Press.

Zarnowitz, V., and Braun, P. 1989. Comment [on Stock and Watson 1989]. In O. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual—1989,* pp. 397–408. Cambridge, Mass.: MIT Press.

Zarnowitz, V., and Moore, G. 1981. The timing and severity of the recession of 1980. *NBER Reporter* (Spring), pp. 19–21.

Zarnowitz, V., and Moore, G. 1982. Sequential signals of recession and recovery. *Journal of Business* 55 (January): 57–85. Revised in G. Moore (ed.), *Business Cycles, Inflation, and Forecasting*. Cambridge, Mass.: Ballinger, 1983.

Zarnowitz, V., and Moore, G. 1986. The development and role of the National Bureau of Economic Research's business cycle chronologies: In R. Gordon (ed.), *The American Business Cycle: Continuity and Change*. Chicago: University of Chicago Press.

Zarnowitz, V., and Moore, G. 1991. Forecasting recessions under the Gramm-Rudman-Hollings law. In K. Lahiri and G. Moore (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records*. Cambridge: Cambridge University Press.

Zerwitz, D. 1989. Business cycles. *NBER Reporter* (Spring), pp. 1–3.