Journal of Econometrics 178 (2014) 368–381

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Estimating turning points using large data sets

James H. Stock [a,*], Mark W. Watson [b]

[a] Department of Economics, Harvard University and the National Bureau of Economic Research, United States

[b] Department of Economics and the Woodrow Wilson School, Princeton University and the National Bureau of Economic Research, United States

ARTICLE INFO

Article history:
Available online 20 September 2013

ABSTRACT

Dating business cycles entails ascertaining economy-wide turning points. Broadly speaking, there are two approaches in the literature. The first approach, which dates to Burns and Mitchell (1946), is to identify turning points individually in a large number of series, then to look for a common date that could be called an aggregate turning point. The second approach, which has been the focus of more recent academic and applied work, is to look for turning points in a few, or just one, aggregate. This paper examines these two approaches to the identification of turning points. We provide a nonparametric definition of a turning point (an estimand) based on a population of time series. This leads to estimators of turning points, sampling distributions, and standard errors for turning points based on a sample of series. We consider both simple random sampling and stratified sampling. The empirical part of the analysis is based on a data set of 270 disaggregated monthly real economic time series for the US, 1959–2010.

1. Introduction

The determination of business cycle turning points is a classic problem in economic statistics. Many of our basic notions of the lead–lag relations among macroeconomic time series are informed by traditional methods of dating turning points for individual series and comparing them to turning points of the overall economy. Chronologies of business cycle turning points (in the jargon, reference cycle chronologies) are currently maintained in the United States by the NBER Business Cycle Dating Committee, in Europe by the CEPR, and by similar organizations in other countries.

This paper compares two approaches to dating business cycles. The dominant current approach, both in the academic literature and in the real-time practice of dating committees, is to date reference cycles by focusing on one, or a few, highly aggregated time series. Hamilton (2011) surveys the academic literature on identifying peaks (dating and predicting recessions). All the methods he discusses define recessions or turning points in terms of single highly aggregated series such as GDP or a monthly index of coincident indicators. Press releases of the NBER Business Cycle Dating Committee indicate that its current practice is to focus on a few highly aggregated series; for example, the press release announcing the 2007:12 peak (NBER (2008)) gives greatest weight to three aggregates (establishment employment, GDP, and GDI), gives secondary weight to five more aggregates (industrial production, household employment, real manufacturing and trade sales, real personal income less transfers, and monthly consumption), and mentions no other series. We will use the term "average then date" to describe the dating of reference cycles using a single highly aggregated series, such as GDP.

As Harding and Pagan (2006) point out, this average-then-date approach contrasts with the approach of the pioneers of business cycle dating, who considered a large number of disparate disaggregated series, identified turning points in those disaggregated series, then determined reference cycle turning points based on the distribution of the turning points of the disaggregated series; see Burns and Mitchell (1946, p. 13 and pp. 77–80). We refer to this latter approach as "date then average".

This paper makes six contributions to the literature on dating reference cycles. The first is to specify a nonparametric estimand which constitutes a population definition of a turning point. The estimand we focus on is the local mode of the population distribution of turning points of disaggregated coincident economic indicators, although we consider other local measures of central tendency as well.[1] This nonparametric population definition of an

* Corresponding author.
E-mail address: james_stock@harvard.edu (J.H. Stock).

0304-4076/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jeconom.2013.08.034

[1] An alternative would be to consider the mode if a clear mode exists or, if not, the end of a plateau in the population distribution of turning points. This alternative is consistent with Burns and Mitchell (1946, pp. 77–80): "In many cases the turning points of different series were bunched so closely that we could not go far astray. But there were cases in which the turning points were widely scattered, and others in which they were concentrated around two separate dates. If there was little else to guide us, we placed the reference turn toward the close of the transition period". It is not clear how to formulate this "close of transition period" scheme mathematically so we restrict attention to local measures of central tendency, with primary focus on the mode. See Harding and Pagan (2006, Section 4) for additional discussion of the role of clusters in Burns and Mitchell (1946).

estimand contrasts with methods in which turning points are defined within a parametric model (e.g. Hamilton (1989)), are defined by an algorithm applied to a realization of series (e.g. Harding and Pagan (2006), in which a turning point is a sample, not population, concept), or are based on expert judgment. Second, with an estimand in hand we undertake statistical inference for date-then-average business cycle turning points. For example, we use asymptotic theory for the kernel estimator of the mode to compute confidence intervals for reference cycle dates estimated from turning points of a random sample of disaggregated series. Third, in the data set we use, some series are available for only a subset of the period and the series are sampled such that different high-level aggregates are not equally represented, and the possibility arises that these departures from simple random sampling could introduce bias by overweighting certain leading or lagging series. We therefore provide and implement methods for adjusting for these sampling irregularities using a stratified sampling framework. Fourth, the empirical work uses a large number of disaggregated series; specifically, the data set consists of 270 monthly real economic activity indicators for the United States, 1959:1–2010:9, where the series are components of four categories of indicators: employment, industrial production, personal income, or sales. Fifth, we adapt some graphical tools (enhanced heat maps) which convey the turning point features of the data. Sixth, this paper also makes a contribution to the average-then-date literature by considering chronologies based on three new monthly measures of GDP developed in Stock and Watson (2010a): an expenditure-based monthly GDP (MGDP-E), an income-based monthly GDP (MGDP-I), and their geometric average (MGDP).

The date-then-average definition of the reference cycle depends on the population of economic series under consideration. The empirical work in this paper uses disaggregates of series that have long served as roughly coincident monthly measures of economic activity: production, sales, income, and employment. An ongoing debate is whether to define business cycles in terms of output, employment, and both; here we take the more comprehensive view but the methods developed here apply equally to narrower definitions.

There is a fairly large literature on business cycle dating using modern time series methods, recently surveyed by Hamilton (2011). The papers most closely related to this one are Harding and Pagan (2006), Chauvet and Piger (2008), and Stock and Watson (2010b). Harding and Pagan (2006) is the first modern paper we are aware of to attempt to formulate the Burns and Mitchell (1946) approach of establishing reference cycle turning points from turning points of multiple individual series. Chauvet and Piger (2008) implement the Harding and Pagan (2006) approach in real time and compare it with an average-then-date chronology based on a Hamilton (1989) Markov switching filter. Both Harding and Pagan (2006) and Chauvet and Piger (2008) consider a small number of series (four), and neither provide a statement of the estimand or standard errors. Stock and Watson (2010b) present preliminary results on reference cycle turning points estimated using the 270-series data set. Their turning points are computed as unadjusted means of individual-series turning points; the results provided here improve upon Stock and Watson (2010b) by estimating in addition the mode and the median and by adjusting for data irregularities.

The date-then-average methods are described in Section 2 and the average-the-date methods are described in Section 3. The data set and the empirical results are presented in Section 4, and Section 5 concludes.

2. Date-then-average methods for reference cycle dating

We consider the problem of dating a reference cycle turning point (peak or trough), conditional on the event that a single turning point occurred in a given episode covering a known time span. This corresponds to a situation in which it is known that a recession occurred during a particular time interval and all that remains is to date the peak within this interval. Conditioning on an episode known to contain a peak or trough is done as an analytical simplification. An extension of the methods here would be to examine estimators that determine simultaneously whether there is a recession and the date of the recession (as in the Harding and Pagan (2006) algorithm).

This section first describes date-then-average reference cycle dating with a simple random sample of series. In our data set, sampling is better thought of as stratified sampling with unequal weights and long periods of missing data, and we propose two modifications of the methods for simple random sampling to handle these data irregularities. Throughout this paper, turning points for individual series are calculated using the Bry and Boschan (1971) algorithm.[2]

2.1. Dating using a simple random sample of disaggregated series

We imagine a population of economic time series, each of which measures a different aspect of economic activity; we approximate this population as being infinitely large. In general, a member of this population has turning points. Thus in a given episode s, which covers a known time interval, there exists a population distribution of dates of turning point of specific series in the population. Letting τ denote the turning point of an individual series, we denote this population distribution of turning points as $g_s(\tau)$. The estimand, that is, the reference cycle turning point, is defined as a functional of this distribution. For reasons discussed in the introduction, we focus on the mode, which we denote D_s^{mode}, however we also consider the median (D_s^{med}) and the mean (D_s^{mean}).

The mean, median, and mode of the distribution g_s can be estimated from a sample of turning points, $\{\tau_{is}\}, i = 1, \ldots, n_s$ where τ_{is} is the turning point date of series i in episode s and n_s is the number of turning points observed in episode s. Let $\hat{D}_s^{\text{mean}}$, $\hat{D}_s^{\text{med}}$, and $\hat{D}_s^{\text{mode}}$ respectively denote the sample mean, median, and mode computed using the sample $\{\tau_{is}\}$. We compute the mode as the mode of a kernel density estimator $\hat{g}_s$ of g_s, with kernel K and bandwidth h.

If the sample of series is obtained by simple random sampling from the population of series then the turning points are *i.i.d.* and the asymptotic distributions of the three estimators are,

$$\sqrt{n_s}(\hat{D}_s^{\text{mean}} - D_s^{\text{mean}}) \xrightarrow{d} N(0, \sigma_{\tau,s}^2), \tag{1}$$

$$\sqrt{n_s}(\hat{D}_s^{\text{med}} - D_s^{\text{med}}) \xrightarrow{d} N\left(0, \frac{1}{4g_s(D_s^{\text{med}})^2}\right), \quad \text{and} \tag{2}$$

$$\sqrt{n_s h^3}(\hat{D}_s^{\text{mode}} - D_s^{\text{mode}}) \xrightarrow{d} N\left(0, \frac{g_s(D_s^{\text{mode}}) \int \left[K'(z)\right]^2 dz}{\left[g_s''(D_s^{\text{mode}})\right]^2}\right), \tag{3}$$

where $\sigma_{\tau,s}^2 = \text{var}(\tau_{is})$ in episode s. Result (3) for the estimator of the mode dates to Parzen (1962), also see Romano (1988) and

[2] The first step of the Bry–Boschan algorithm entails a nearly-centered 15-month moving average. We found that this occasionally produced some anomalous results, specifically peaks lower than their counterpart troughs, and that these anomalies were eliminated by using a centered 3-month moving average. The results reported in this paper therefore all use the three-month moving average in the first Bry–Boschan step.

Ziegler (2003); for this result the bandwidth sequence h_n satisfies $h_n \to 0$, $nh_n^3 \to \infty$. The variances in (1)–(3) are consistently estimable using kernel estimators of g_s and its second derivative, g_s''.

2.2. Adjusting for weighted random sampling: lag adjustment

As is discussed in Section 4, in our data set components of industrial production are more heavily represented than components of personal income, and the relative number of components varies over time. If turning points in industrial production systematically lead turning points in personal income then treating our sample of turning points as a simple random sample will bias the estimator toward an estimated reference cycle turning point that leads the population reference cycle turning point.

We model this problem of unequal representation of classes of series as one of stratified sampling, in which the initial stratum is the class of series (such as industrial production). The subaggregate (such as industrial production of primary metals) is then randomly sampled within the class. The number of observations (series) differs from one class to the next. This results in some classes of series receiving larger weight in the sample than in the population. Bias arises if turning points in a class of series lead or lag the population reference cycle turning point and if the sample and population weights differ by series class.

We use two different procedures for adjusting for discrepancies between sample and population weights by series class: lag adjustment and weighted estimation. Both procedures involve weighting by the ratio of population to sample probabilities. Let m index the classes of series, M the number of classes (which we take to be finite), m_i the class containing series i, π_m the population probability assigned to class m, and p_{ms} the fraction of series of class m in the sample of turning points for episode s. Then the ratio w_{is} of population weights to sample weights for series i in episode s is,

$$w_{is} = \frac{\pi_{m_i}}{p_{m_i s}}. \tag{4}$$

Lag adjustment by class. The first procedure exploits the panel nature of the data set by estimating a mean lag for each series. Let series in class m have a population mean lag k_m, relative to the reference cycle date Then we can write the turning point of the ith series in episode s as the sum of the population mean reference cycle turning point D_s^{mean}, the mean lag for its class, and a discrepancy η_{is}:

$$\tau_{is} = D_s^{\text{mean}} + k_{m_i} + \eta_{is}. \tag{5}$$

The reference cycle turning point is identified as the mean by assuming that $E\eta_{is} = 0$ and that k_m are normalized so that the mean lag in population is zero. This latter condition corresponds to $\sum_{m=1}^{M} \pi_m k_m = 0$.

Unless $w_{is} = 1$ for all i and s, in general estimation of (5) by OLS will not satisfy the restriction $\sum_{m=1}^{M} \pi_m k_m = 0$ and the estimates of the class lags will be biased. The lag adjustment procedure therefore has two steps. First, $\{k_m\}$ in (5) are estimated by restricted least squares, subject to the restriction that $\sum_{m=1}^{M} \pi_m k_m = 0$; this yields the estimators $\{\hat{k}_m\}$. Second, the sample of adjusted turning points is constructed as $\tilde{\tau}_{is} = \tau_{is} - \hat{k}_{m_i}$. The mean, median, and mode estimators are then computed episode-by-episode using the lag-adjusted data, $\{\tilde{\tau}_{is}\}$.

2.3. Adjusting for weighted random sampling: weighted estimation

The second procedure for adjusting for weighted random sampling involves weighting the sample so that the sample weights on individual observations (that is, series-specific turning points) match the population weights.

Let g_{ms} denote the distribution of turning points among series of class m in episode s. The population distribution of turning points in episode s is then $g_s = \sum_m \pi_m g_{ms}$, where the sum is over the finitely many classes of series. Because the Bry–Boschan algorithm produces integer-valued turning points, the raw data consist of histograms of turning point dates for each class of series, by episode. The weighted estimation schemes are all based on weighting these histograms of turning points by class to yield a weighted histogram, where the weights are the ratio of the population to sample weights. Specifically, the weighted histogram for episode s is $\hat{g}_s^{hist-wtd}(t) = \sum_{i=1}^{n_s} I(\tau_{is} = t) / \sum_{i=1}^{n_s} w_{is}$. The weighted mean and median are computed directly from the weighted histogram. The weighted mode is computed as the mode of the kernel density estimator computed by smoothing $\hat{g}_s^{hist-wtd}$.

Variances for the weighted mean and median estimators are,

$$\text{var}(\hat{D}_s^{\text{mean,wtd}}) = \sum_m \left(\frac{\pi_m^2}{n_{ms}}\right) \sigma_{ms}^2 \tag{6}$$

$$\begin{aligned} \text{var}(\hat{D}_s^{\text{med,wtd}}) &= \sum_m \left(\frac{\pi_m^2}{n_{ms}}\right) \frac{G_{ms}(D_s^{\text{med}})\left[1 - G_{ms}(D_s^{\text{med}})\right]}{g_s(D_s^{\text{med}})^2} \\ &\le \frac{1}{n_s}\left(\frac{1}{4g_s(D_s^{\text{med}})^2}\right) \sum_m \frac{\pi_m^2}{p_m}, \end{aligned} \tag{7}$$

where G_{ms} is the cdf corresponding to g_{ms}.

Because the terms in the first summation in (7) are poorly estimated, the empirical work uses the bound in the final expression in (7) for the standard errors for the weighted median.

The variance of the weighted mode is that given in (3) for the mode under simple random sampling, with the modification that g_s is reinterpreted as the weighed density. Standard errors for the mode of the weighted distribution are computed using the kernel smoother of the weighted histogram to estimate g_s.[3]

3. Average-then-date methods

Dating using aggregates entails identifying turning points (here, using the Bry–Boschan algorithm) in an aggregate measure of economic activity. We consider six such measures. Three of these are indexes of coincident economic indicators, constructed as weighted average of four monthly aggregates: industrial production (IP), nonfarm employment (EMP), real manufacturing and wholesale–retail trade sales (MT), and real personal income less transfers (PIX). The remaining three measures of aggregate activity are monthly estimates of quarterly GDP. We use data on the aggregates series from 1959:1–2010:6.

3.1. Indexes of coincident economic indicators

Let X_{it} denote one of the four series (IP, EMP, MT, and PIX) in native units in period t, so $i = 1, \ldots, 4$, and let $y_{it} = \Delta\ln(X_{it})$. We consider three indexes C_t constructed from these series.

1. The index published monthly by The Conference Board (TCB), which is normalized to equal 100 in 2004:7.

[3] The idea of stratified sampling could be carried further than we do here by including additional substrata. For example, below manufacturing and trade sales there exists an industry stratum, e.g. durables manufacturing, nondurables manufacturing, etc. We have assumed independence across turning points within our single-stratum sampling unit (e.g. among components of manufacturing and trade sales). One could relax this by allowing for clustering at a lower stratum. This extension to clustered standard errors is left to future work.

Table 1
Series weights for the coincident indexes.

Series	Coincident index		
	CI-TCB	CI-ISD	CI-DFM
IP	0.13	0.14	0.58
EMP	0.50	0.49	0.07
MT	0.11	0.11	0.23
PIX	0.26	0.26	0.12

Notes: Weights for the TCB index were estimated by a regression of the change in the index on the change in the four components ($R^2 = 0.97$). Weights for the dynamic factor model (DFM) coincident index are the sum of the steady-state Kalman smoother weights on current, lead, and lagged values of the row series.

2. An index constructed by inverse standard deviation weighting (ISD): $C_{it}^{ISD} = \exp\left[\sum_{i=1}^{4} \alpha_i \ln(X_{it})\right]$, where $\alpha_i = s_i^{-1} / \sum_{j=1}^{4} s_j^{-1}$ and s_i is the (full-sample) standard deviation of y_{it}. The index is normalized to equal 100 in 2004:7.
3. An index constructed as the estimated common factor from a dynamic factor model of the four variables, with a single factor (DFM). The model is similar to that in Stock and Watson (1989), with different lag specification. The model used here is,

$$y_{it} = \alpha_i + \lambda_i f_t + u_{it}$$
$$f_t = \alpha_f + \phi_i f_{t-1} + \phi_2 f_{t-2} + \varepsilon_t$$
$$u_{it} = \rho_{i1} u_{it-1} + \rho_{i2} u_{it-2} + e_{it}$$
$$\begin{bmatrix} \varepsilon_t \\ e_{1t} \\ e_{2t} \\ e_{3t} \\ e_{4t} \end{bmatrix} \sim \text{i.i.d. } N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\varepsilon & 0 & 0 & 0 & 0 \\ 0 & \sigma_1 & 0 & 0 & 0 \\ 0 & 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & 0 & \sigma_4 \end{bmatrix} \right).$$

The parameters are estimated by Gaussian maximum likelihood, using data that have been adjusted for outliers. Let the steady-state Kalman smoother estimator of f be $f_{t/T} = \beta_0 + \sum_{i=1}^{4} \sum_{k=-\infty}^{\infty} \beta_{ij} y_{it-k}$; the model parameters are normalized so that $\beta_0 = 0$ and $\sum_{i=1}^{4} \sum_{k=-\infty}^{\infty} \beta_{ij} = 1$. Let $\hat{f}_{t/T}$ denote smoothed values computed using the estimated parameters, applied to y_{it} computed using the original data (not outlier-adjusted). The DFM coincident index is $C_t^{DFM} = \exp(\hat{f}_{t/T})$, which is then scaled to equal 100 in 2004:7.

Table 1 gives the weights for the three indexes. These indexes have quite different implied weights. The ISD index puts nearly half the weight on employment and very little on IP and MT. In contrast, the DFM index places over half the weight on IP and very little on EMP. The weights for the TCB index are quite close to the inverse standard deviation weights. It is important to note that one reason the chronologies based on these different indexes differ is that by weighting the different series differently, the indexes have different average growth rates and different standard deviations, which leads to different periods of negative growth.

3.2. Monthly GDP

We also consider dates based on three monthly measures of real GDP. Construction of these measures is described in Stock and Watson (2010a). Briefly, there are two separate measures, an expenditure-based monthly GDP which, following Nalewaik (2010) we refer to as GDP(E), and an income-based monthly GDP, which we refer to as GDP(I). Nominal monthly GDP(E) is estimated as the sum of eight components (consumption, investment in nonresidential structures, investment in residential structures, investment in equipment and software, change in inventories, exports, imports, and government purchases). Some of these are observed on a monthly basis. For components that are only reported quarterly, the quarterly values are distributed using a state space model

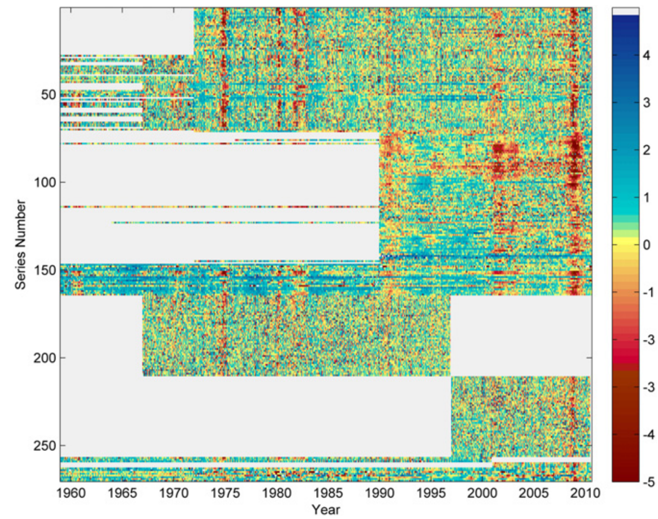


Fig. 1. Heat map of monthly growth rates divided by the series standard deviation for the 270 series in the monthly data set. The vertical axis is the series number as given in the Appendix A; the horizontal axis is the monthly time scale, 1959:1–2010:9. Negative monthly growth appears as red, positive monthly growth rates appear as blue. The gray sections indicate missing data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in which the monthly concept is modeled as a latent series that is correlated with observable monthly series chosen to be conceptually close to the specific component. Real monthly GDP(E) is computed from the nominal series using a monthly interpolation of the GDP price deflator. Nominal and real monthly GDP(I) are computed analogously, using six components (employee compensation, proprietors' income, rental income, net interest, corporate profits, and other).

We also use an estimate of monthly GDP that combines the expenditure- and income-based estimates. This combined series, GDP(Avg), is computed as the geometric average of GDP(E) and GDP(I).

4. Empirical results

4.1. The disaggregated data set

The disaggregated data set consists of 270 components of industrial production (69 distinct component series), nonfarm employment (95 series), real manufacturing and wholesale–retail trade sales (92 series), and real personal income less transfers (14 series). All data are monthly for the United States, with a maximum span of 1959:1–2010:9 (621 months). The series and the spans for which they are available are listed in the Appendix A.

The monthly growth rates of the 270 series, divided by their standard deviation, is displayed as a heat chart in Fig. 1. (Figs. 1–3, shown in color in the online version of the article, are presented in gray scales as supplemental Figs. S-1–S-3.) The vertical axis is series number as given in the Appendix A, the horizontal axis is time in months. Blue denotes periods of positive growth, yellow denotes moderately negative growth, and red denotes strongly negative growth. The rectangular gray swaths in Fig. 1 represent missing data. The most relevant features of Fig. 1 for the current purpose are the vertical yellow–red bands. Because the horizontal axis is calendar time, the vertical yellow–red bands show periods in which many of the component series were experiencing negative growth. In the context of Fig. 1, the task is to date the beginning and end of the yellow–red band, which are respectively the cyclical peak and trough.

The periods of negative growth are more apparent in Fig. 2, which is an enhanced version of Fig. 1. Specifically, Fig. 2 plots

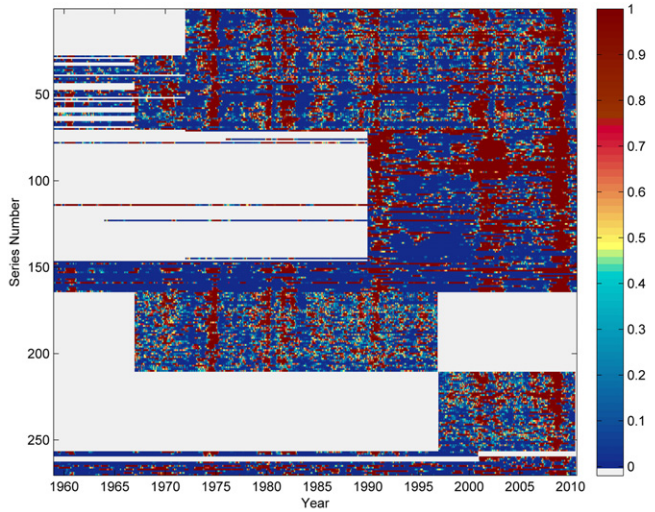


Fig. 2. Enhanced version of the monthly disaggregated data set heat map. The heat map plots $\Phi(-\kappa(z_{it} - \mu))$, where κ and μ are minimum-entropy scale and shift factors and z_{it} is the monthly growth rate of series i divided by its standard deviation.

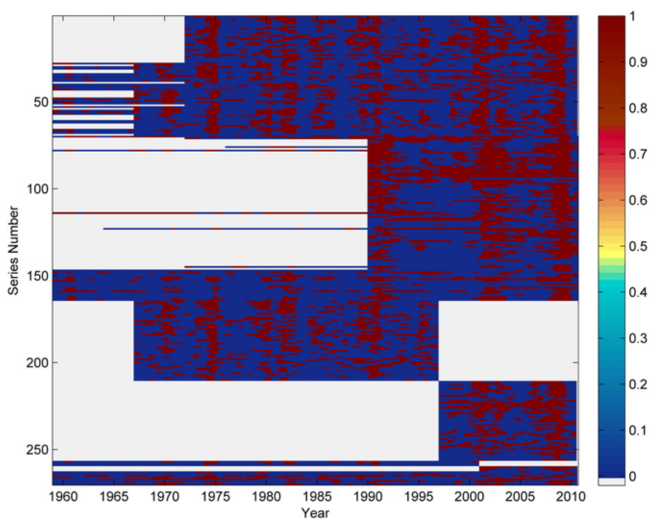


Fig. 3. Bry–Boschan recessions computed using the monthly disaggregated data set. Dark red denotes Bry–Boschan recessions (from a peak to a trough) and blue denotes Bry–Boschan expansions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$\Phi[-\kappa(z_{it} - \mu)]$, where z_{it} growth rate of series i, divided by its standard deviation, Φ is the cumulative normal distribution function, and κ and μ are minimum-entropy scale and location parameters. The NBER-dated postwar recessions clearly stand out as dark red vertical bands. From the perspective of dating turning points, the beginning of each red band suggests a reference cycle peak and the end of the band suggests a trough.

Fig. 3 plots series-specific recession episodes, where a series-specific recession is defined to be the period from a Bry–Boschan peak to a Bry–Boschan trough. The NBER recessions remain clearly visible in Fig. 3. It is evident that there is considerable dispersion of specific-series turning points around the beginning and end of the recessions. In addition, there are evidently many more or less random specific-series recessions that do not align with recessions in other series.

As discussed above, our analysis focuses on dating turning points, conditional on a turning point having occurred. Figs. 1–3 suggest that the disaggregated series would also be useful for ascertaining whether a turning point has occurred – that is, determining the presence of a vertical band – but we do not pursue that. Henceforth, we focus on data by episode, where an episode is defined to be the NBER turning point date ±12 months.

4.2. Results: reference cycle chronologies

We begin with reference cycle chronologies computed using the aggregate monthly series, then turn to chronologies based on the distribution of turning points in the disaggregated series.

Average-then-date chronologies. Fig. 4 plots two of the monthly aggregate series, the inverse standard deviation-weighted coincident index (CI-ISD) and the combined monthly GDP series GDP(Avg). The vertical lines in Fig. 4 represent the reference cycle chronologies based on the plotted series and the NBER chronology. The coincident index is more cyclically volatile than the monthly GDP series. One notable difference between the two series is that monthly GDP plateaus during 2001 but the Bry–Boschan algorithm does not indicate a recession; this is not surprising because quarterly (expenditure-based) GDP declines for only one quarter, and grows over every two-quarter period, during this episode. In most episodes, however, turning points based on the two series coincide and also coincide with the NBER chronology.

The chronologies based on all six monthly aggregates (the three coincident indexes and the three monthly GDP series) are summarized in Table 2, as leads or lags relative to the NBER date. The coincident indexes produce the same turning points in 9 of the 16 cases, and are within a month of each other in all but 4 cases. Notably, the DFM index, which puts considerable weight on industrial production, dates the 1969:12 and 1980:1 peaks earlier than the other coincident indexes, and the TCB index dates the 2001:11 trough four months later. Of these three indexes, the CI-ISD comes the closest to matching the NBER chronology, with a mean absolute difference between its chronology and the NBER chronology of 0.80 months; the only discrepancy between the CI-ISD and NBER chronologies exceeding two months is the 2001:3 peak.

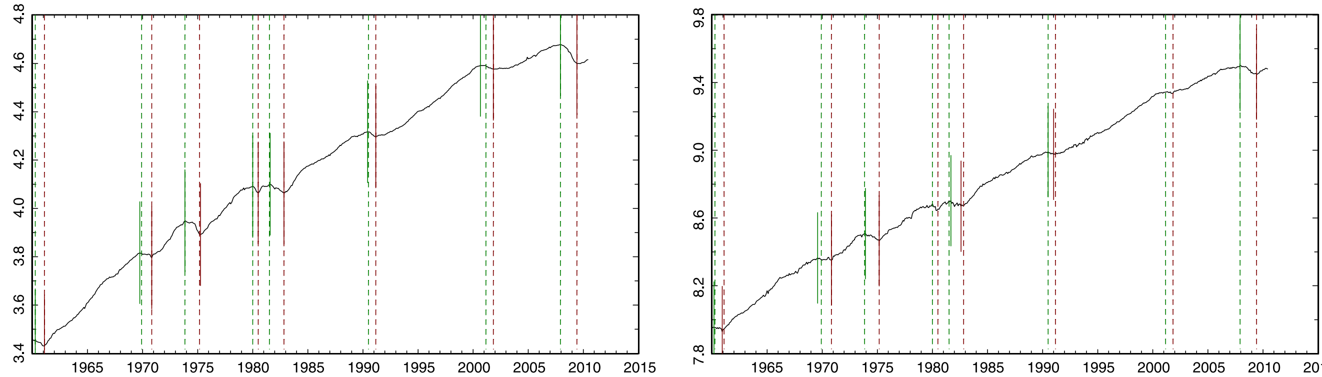
The monthly GDP(E) chronology differs substantially from the NBER chronology and the coincident index chronologies: it dates 5 turning points earlier than the corresponding NBER date, by as much as 10 months, and does not identify a recession in 1980 or 2001.[4] The GDP(I) chronology does not identify a recession in 1970 and dates the 2007:12 peak 12 months earlier, but otherwise is close to the NBER chronology. The final column of the table displays the chronology for the Macroeconomic Advisors monthly GDP series, which is available only since 1992:4. Like GDP(E), this series does not detect a recession in 2001 and dates the 2007:12 peak six months earlier than the NBER.

Date-then-average chronologies. Table 3 reports three sets of date-then-average chronologies based on the 270-series disaggregated data set, along with their standard errors.[5] The first set of chronologies, reported in the first three columns of results, are unadjusted turning point estimates, for which the series are treated as if they resulted from simple random sampling (Section 2.1). The next block of three columns reports chronologies based on the weighted lag adjustment procedure described in Section 2.2, with class-specific fixed effects as in (5).[6] The final block of columns reports the results of weighted estimation, computed following Section 2.3. The lag-adjusted and weighted estimation methods require population weights $\pi_1, \ldots, \pi_4$ for the four classes of series. The results in Table 3 use population weights of 0.3 for IP,

4 GDP(E) declined sharply in 1980, however the decline only lasted 5 months so it is not identified as a recession by the Bry–Boschan algorithm.

5 The kernel density estimator $\hat{f}_s$ was computed using the biweight kernel, $K(z) = (15/16)(1-z^2)^2$, for which $\int \left[K'(z)\right]^2 dz = 2.1429$, with bandwidth $h = 4$ months.

6 The class-specific estimated lags (k_m in the notation of (5)) are -0.82 for IP, 2.65 for EMP, -1.12 for MT, and -2.01 for PIX.



(a) Monthly coincident index, inverse standard deviation weighting.

(b) Monthly GDP(Avg).

Fig. 4. The inverse standard deviation-weighted monthly coincident index (panel a), monthly GDP(Avg) (panel b), Bry–Boschan turning points for each series (solid vertical lines), and NBER chronologies (dashed vertical lines). Peaks are green, and troughs are red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Average-then-date chronologies computed using three monthly coincident indexes and four measures of monthly GDP, as a lead (positive value) or lag (negative value) of the NBER turning point.

NBER		Coincident indexes			Monthly GDP			
		CI-TCB	CI-ISD	CI-DFM	GDP(E)	GDP(I)	GDP(Avg)	GDP-MA
1960:4	P	−2	0	–	−1	−2	−1	
1961:2	T	0	0	0	−2	−2	−2	
1969:12	P	−2	−2	−4	−4	–	−4	
1970:11	T	0	0	0	−10	–	0	
1973:11	P	0	0	0	1	0	1	
1975:3	T	1	1	1	0	−1	0	
1980:1	P	0	0	−10	–	0	–	
1980:7	T	0	0	0	–	−1	–	
1981:7	P	0	1	0	2	1	2	
1982:11	T	0	0	0	−6	0	−3	
1990:7	P	−1	−1	0	0	0	0	
1991:3	T	0	0	0	0	−2	−2	
2001:3	P	−6	−6	−6	–	0	–	–
2001:11	T	4	0	0	–	−1	–	–
2007:12	P	−1	0	0	1	−12	0	1
2009:6	T	0	0	0	0	1	0	0
Mean		−0.44	−0.44	−1.27	−1.58	−1.36	−0.75	0.50
MAE		1.06	0.69	1.40	2.25	1.64	1.25	0.50

Notes: Entries are the NBER turning point minus the series-specific Bry–Boschan turning point, in months. Episodes for which the series is available but does not have a Bry–Boschan turning point are denoted by "–". The GDP(E), GDP(I), and GDP(Avg) monthly GDP series are from Stock and Watson (2010a). The GDP-MA series is the Macroeconomic Advisors Monthly GDP series, which starts in 1992:4. The mean and mean absolute error (MAE) in the final two rows summarize the discrepancies of the chronology for the column series, relative to the NBER chronology; episodes in which a series does not have a Bry–Boschan recession are excluded from the summary statistics.

EMP, and MT, and 0.1 for PIX. This choice of unequal weights was made for the practical reason that in some episodes there are very few (as few as 2) PIX subaggregate turning points so the weighting scheme of Section 2.3 results in those series getting very large weights and produces some outliers. The large weights and outliers call into question the validity of the asymptotic standard errors. Results based on equal population weights for the four classes are discussed below as a sensitivity check.

The date-then-average chronologies in Table 3 have three noteworthy features. First, the standard errors of the estimated turning points are fairly small, in most cases in the range 0.5 to 0.8, so that a typical confidence interval for a turning point is ± 1 to ± 1.6 months. The standard errors tend to be larger for earlier episodes, which is consistent with the number of series increasing over the course of the sample.

Second, even though the sample does not have equal representation of the classes of series, using the adjusted methods (lag adjustment and weighted estimation) makes surprisingly little difference to the estimated chronology. For example, the greatest difference between the mode chronology using no adjustments, compared with the weighted mode, is 0.5 months.

Third, the mean, median, and mode estimates typically agree rather closely when computed using the same adjustment procedure (unadjusted, lag-adjusted, or weighted), although there are several episodes in which the three estimators differ by up to two months. In this sense, choice of estimand matters for the resulting chronology, at least in some episodes.

Fourth, in most episodes the date-then-average chronologies are very close to the NBER chronology, but there are a few episodes with notable differences. For example, the weighted mode chronology differs from the NBER chronology by less than one month in 12 of the 16 episodes. However, all the date-then-average chronologies date the 1969:12 and 2001:3 peaks earlier than the NBER, and for these peaks the difference between the date-then-average turning point and the NBER turning point is statistically significant at the 5% level for nearly all estimators in the table.

Sensitivity analysis. We briefly discuss two other sets of results as sensitivity checks. First, Table 2 was also computed using equal population weights, that is, $\pi_1 = \cdots = \pi_4 = 0.25$. This affects only the second two blocks of results (lag-adjusted and weighted estimation). Using equal population weights produced changes for the class-lag adjusted estimates of typically 0.1 months or less.

Table 3
Date-then-average chronologies and standard errors computed using turning points of 270 disaggregated series, as a lead (positive value) or lag (negative value) of the NBER turning point.

NBER Dates		No adjustments			Class lag-adjusted			Weighted estimation		
		Mean	Median	Mode	Mean	Median	Mode	Mean	Median	Mode
1960: 4	P	−1.8 (0.6)	−2.0 (0.7)	−1.4 (0.5)	−2.5 (0.7)	−2.3 (0.8)	−2.5 (0.3)	−2.0 (0.6)	−2.0 (0.3)	−1.4 (0.4)
1961: 2	T	−0.3 (0.4)	0.0 (0.6)	−0.5 (0.7)	−0.8 (0.3)	−1.1 (0.5)	−0.5 (0.2)	−0.3 (0.3)	0.0 (0.3)	−0.6 (0.2)
1969:12	P	−2.2 (0.7)	−2.0 (0.6)	−2.3 (0.4)	−1.7 (0.6)	−1.8 (0.7)	−1.3 (0.5)	−1.7 (0.8)	−2.0 (0.4)	−2.4 (5.9)
1970:11	T	1.2 (0.6)	0.0 (0.7)	−0.2 (0.4)	1.7 (0.6)	1.2 (0.7)	0.7 (0.3)	1.9 (0.7)	1.0 (0.6)	0.1 (2.7)
1973:11	P	1.3 (0.6)	2.0 (0.6)	1.6 (0.3)	1.9 (0.6)	3.0 (0.7)	2.2 (0.3)	2.4 (0.7)	3.0 (0.7)	1.7 (1.0)
1975: 3	T	1.0 (0.3)	0.0 (0.3)	0.4 (0.3)	1.6 (0.3)	1.2 (0.3)	1.0 (0.1)	1.3 (0.3)	1.0 (0.4)	0.8 (0.8)
1980: 1	P	−1.8 (0.7)	−1.0 (0.8)	−0.3 (0.4)	−1.3 (0.7)	−1.2 (0.9)	0.3 (0.2)	−1.8 (0.9)	−2.0 (0.8)	−0.1 (0.2)
1980: 7	T	−0.9 (0.5)	0.0 (0.4)	−0.5 (0.2)	−0.1 (0.5)	0.2 (0.3)	0.3 (0.1)	−0.5 (0.7)	0.0 (0.4)	0.0 (0.2)
1981: 7	P	−0.7 (0.5)	0.0 (0.5)	−0.1 (0.3)	−0.2 (0.5)	0.2 (0.5)	0.5 (0.1)	−0.1 (0.5)	0.0 (0.4)	0.1 (4.4)
1982:11	T	−0.6 (0.6)	0.0 (0.6)	1.1 (0.4)	−0.2 (0.6)	0.9 (0.6)	1.9 (0.2)	−0.5 (0.6)	0.0 (0.5)	0.9 (0.9)
1990: 7	P	−0.8 (0.6)	0.0 (0.7)	0.3 (0.5)	−0.3 (0.6)	−1.2 (0.8)	1.8 (0.4)	−1.1 (0.6)	−1.0 (0.5)	−0.3 (0.2)
1991: 3	T	2.1 (0.5)	1.0 (0.4)	0.4 (0.3)	2.1 (0.4)	1.1 (0.4)	0.4 (0.1)	2.0 (0.4)	1.0 (0.4)	0.2 (2.0)
2001: 3	P	−3.7 (0.5)	−3.0 (0.6)	−2.2 (0.3)	−4.1 (0.5)	−4.8 (0.6)	−3.2 (0.2)	−3.7 (0.6)	−3.0 (0.6)	−2.3 (4.4)
2001:11	T	0.2 (0.5)	1.0 (0.5)	0.6 (0.2)	0.5 (0.5)	1.2 (0.5)	1.5 (0.1)	0.6 (0.7)	1.0 (0.7)	0.6 (0.9)
2007:12	P	−1.0 (0.5)	−1.0 (0.9)	−6.1 (0.5)	−1.4 (0.5)	−1.8 (0.7)	−2.8 (0.8)	−1.4 (0.5)	−2.0 (0.9)	−6.0 (1.1)
2009:6	T	1.7 (0.3)	1.0 (0.5)	−0.1 (0.2)	1.5 (0.3)	1.7 (0.4)	1.4 (0.2)	1.6 (0.3)	1.0 (0.5)	−0.2 (0.2)
Mean		−0.39	−0.25	−0.59	−0.20	−0.20	0.10	−0.21	−0.25	−0.56
MAE		1.34	0.88	1.12	1.37	1.56	1.38	1.44	1.25	1.11

Notes: Entries are the NBER turning point minus the date-then-average chronology for that column, in months. Standard errors appear in parentheses. The mean and mean absolute error (MAE) in the final two rows summarize the discrepancies of the chronology for the column series, relative to the NBER chronology.

The weighted mean and median chronologies changed somewhat more, but still by less than 0.5 months. The only substantial change was two dates for the weighted mode, for which the equal weighting scheme produced highly influential outliers.

Second, we also computed lag-adjusted chronologies using series-specific lags instead of class lags, so that (5) has series-specific lags instead of class-specific lags. For most episodes the results using series-specific lags are similar to the results in Table 2, but for some episodes they differ from Table 2 results and the series-specific lag adjusted mean, median, and mode also differ from each other. Overall, the resulting chronologies are outliers, relative to those in Table 2 (including the NBER chronology). We view this as a result of limitations of the data. In our unbalanced panel there are many series that appear in only a few episodes, so series-specific lags are in many cases poorly estimated. Some results based on the counterpart of (5) using series-specific lags, but without weighting (and thus not addressing the bias problem) are presented in Stock and Watson (2010b).

Third, we recomputed the estimates in Table 3 using a wider episode band of 15 months on each side of the NBER date instead of 12 months. (At the time of writing, fewer than 15 months of data are available since the 2009:6 trough so we excluded that trough from this sensitivity check.) The sensitivity of the results to the width of the episode depends on the estimator: the mean is the most sensitive, followed by the median, and the mode is the least sensitive. For example, the weighted mean estimate of the 2001:11 trough, relative to the NBER date, changes from +0.6 months using a 12-month episode band to +2.0 months using a 15-month episode band, whereas the weighted mode only changed from +0.6 months to +0.7 months. This robustness of the mode, and lack of robustness of the mean, to the episode width is not surprising, and this robustness is another virtue of the mode as a turning point estimator.

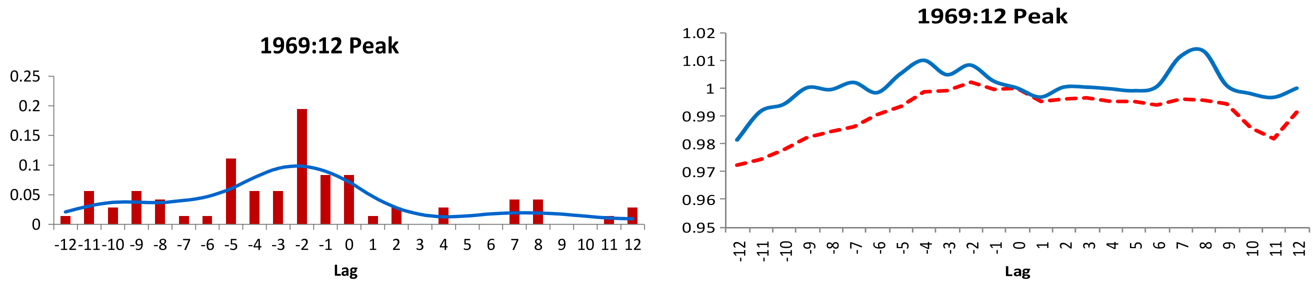
4.3. Four episodes: 1969:12, 1991:3, 2001:3, and 2007:12

We now take a closer look at four episodes in which there is disagreement among the methods examined in Tables 2 and 3.

The 1969:12 peak. The date-then-average methods in Table 3 all date the 1969:12 NBER peak as having occurred between 1.3 and 2.4 months earlier, so that (after rounding) the peak would be 1969:10. Of the average-then-date methods in Table 2, the TCB and ISD chronologies also date 1969:10, whereas the DFM, GDP(E), and GDP(Avg) chronologies place the peak in 1969:8. The GDP(I) series does not detect a Bry–Boschan recession in 1969–70.

Fig. 5a plots the weighted histogram and kernel density estimate of peak turning point dates in the 1969:12 episode, and Fig. 5b plots two monthly aggregates over this episode, monthly GDP(Avg) and the inverse standard deviation-weighted coincident index. The weighted mode of 1969:10 is clearly visible in the kernel density plot and in the weighted histogram. The 1969:8 turning point in GDP(Avg) is due to a local peak that is slightly higher than the 1969:10 local peak. August 1969 is not a local mode in Fig. 5a in either the histogram or the kernel density estimate, nor does it constitute an end-of-transition-period date (rather it is near the beginning of the cluster of turning points from 1969:7 to 1969:12. The date-then-average evidence is consistent with the average-then-date evidence that the peak occurred earlier than 1969:12. Of the two dates suggested by the average-then-date chronologies, the date-then-average evidence points to the later one, October 1969.

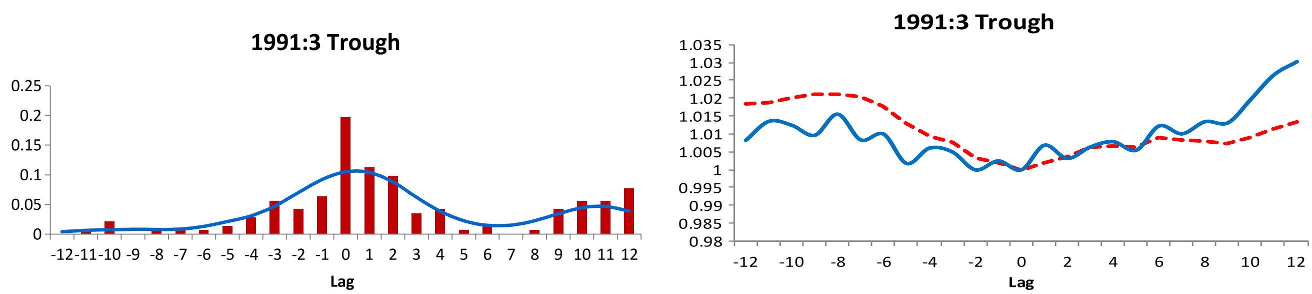
The 1991:3 trough. This episode is interesting because the date-then-average trough estimates based on the means in Table 3 are approximately 2 months later than the median and mode estimates. Moreover, the median and mode estimates are close to the average-then-date troughs, which in all cases except GDP(I) and GDP(Avg) coincide with the NBER trough. The reason for the divergent mean estimate is evident in Fig. 6a. That figure shows two clusters of turning points, the main cluster from 1990:12 to 1991:7, and a smaller cluster in early 1992. The mode and median are in the first cluster which aligns with the average-then-date and NBER chronologies. The mean averages in observations in the second, later cluster and thus produces an estimate that lags the others. Examination of the individual series turning points shows that the smaller cluster is mainly associated with employment series, whose within class mean is 5.4 months for this episode. The discussion in Burns and Mitchell (1946) cited above is consistent with selecting a turning point in the first cluster, which has a distinct mode. The 95% confidence interval for the trough based on the unweighted mode is (1991:2.8, 1991:4.6), which contains 1991:3. In this instance, then, the date-then-average analysis confirms the average-then-date and NBER trough date of 1991:3. This discussion underscores that the mode is preferable to the mean because of the sensitivity of the mean to outliers (distant local turning points for a relatively small number of series).



(a) Weighted histogram and kernel density estimate of turning points of disaggregated series.

(b) Monthly GDP(Avg) (blue solid) and CI-ISD coincident index (red dashed), normalized to 1.00 at the NBER turning point.

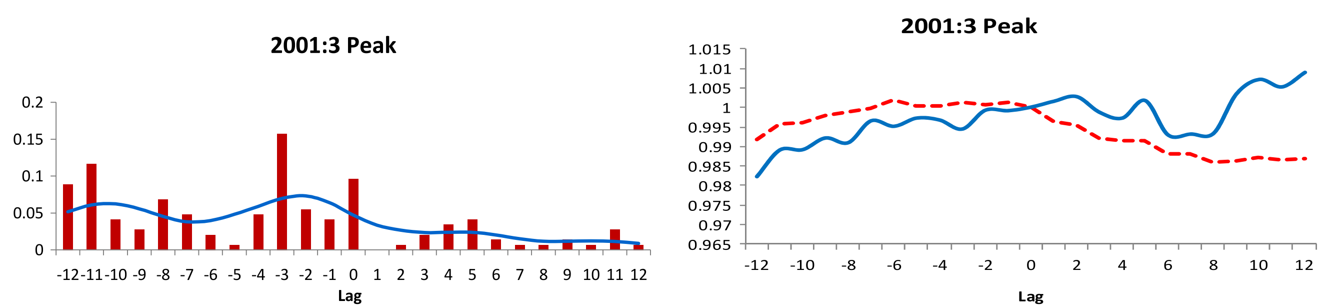
Fig. 5. The 1969:12 NBER peak: (a) date-then-average and (b) average-then-date approaches.



(a) Weighted histogram and kernel density estimate of turning points of disaggregated series.

(b) Monthly GDP(Avg) (blue solid) and CI-ISD coincident index (red dashed), normalized to 1.00 at the NBER turning point.

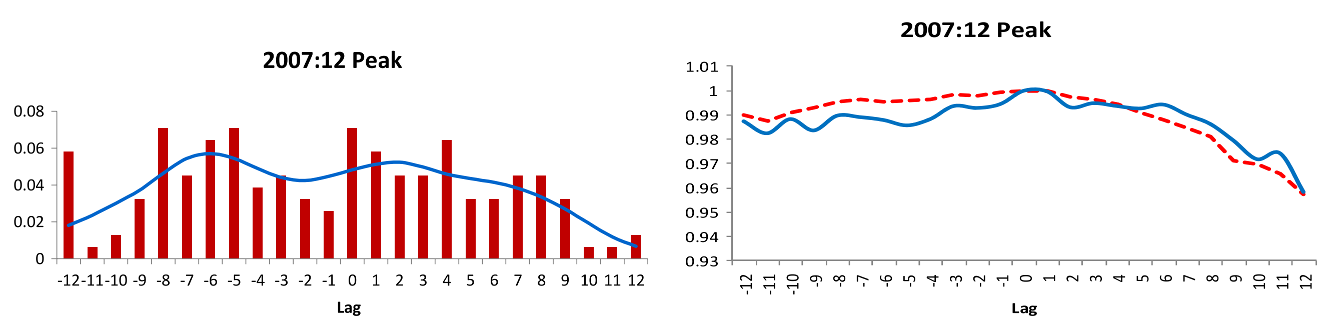
Fig. 6. The 1991:3 NBER trough: (a) date-then-average and (b) average-then-date approaches.



(a) Weighted histogram and kernel density estimate of turning points of disaggregated series.

(b) Monthly GDP(Avg) (blue solid) and CI-ISD coincident index (red dashed), normalized to 1.00 at the NBER turning point.

Fig. 7. The 2001:3 NBER peak: (a) date-then-average and (b) average-then-date approaches.



(a) Weighted histogram and kernel density estimate of turning points of disaggregated series.

(b) Monthly GDP(Avg) (blue solid) and CI-ISD coincident index (red dashed), normalized to 1.00 at the NBER turning point.

Fig. 8. The 2007:12 NBER peak: (a) date-then-average and (b) average-then-date approaches.

The 2001:3 peak. The 2001:3 peak shows considerable date disagreements. The average-then-date estimates based on coincident indexes all estimate 2000:9, the GDP(I) estimate is 2001:3, and the GDP(E) and GDP(Avg) methods do not detect a recession. The date-then average estimates range from 2000:10.2 $\pm$ 1.2 to 2000:12.8 $\pm$ 0.6. The weighted mode estimate is 2000:12.7 with a very wide standard error.

Inspection of Fig. 7a shows two clusters of turning points, one in the spring and summer of 2000 and the second in 2000:11–2001:3, and the kernel density estimate is bimodal. The ISD index is

Table A.1
Series in the dissagregated data set.

No.		Series class	Start	End
	Industrial production by industry			
	Mfg: Durables			
1		IP:Wood product NAICS=321, SA	1972:1	2010:8
2		IP:Nonmetallic mineral product NAICS=327, SA	1972:1	2010:8
3		IP:Primary metal NAICS=331, SA	1972:1	2010:8
4		IP:Fabricated metal product NAICS=332, SA	1972:1	2010:8
5		IP:Machinery NAICS=333, SA	1972:1	2010:8
6		IP:Computer and electronic product NAICS=334, SA	1972:1	2010:8
7		IP:Electrical equipment, appliance, and component NAICS=335, SA	1972:1	2010:8
8		IP:Transportation equipment NAICS=336, SA	1972:1	2010:8
9		IP:Furniture and related product NAICS=337, SA	1972:1	2010:8
10		IP:Miscellaneous NAICS=339, SA	1972:1	2010:8
	Mfg: Nondurables			
11		IP:Food NAICS=311, SA	1972:1	2010:8
12		IP:Beverage NAICS=3121, SA	1972:1	2010:8
13		IP:Tobacco NAICS=3122, SA	1972:1	2010:8
14		IP:Textile mills NAICS=313, SA	1972:1	2010:8
15		IP:Textile product mills NAICS=314, SA	1972:1	2010:8
16		IP:Apparel NAICS=315, SA	1972:1	2010:8
17		IP:Leather and allied product NAICS=316, SA	1972:1	2010:8
18		IP:Paper NAICS=322, SA	1972:1	2010:8
19		IP:Printing and related support activities NAICS=323, SA	1972:1	2010:8
20		IP:Petroleum and coal products NAICS=324, SA	1972:1	2010:8
21		IP:Chemical NAICS=325, SA	1972:1	2010:8
22		IP:Plastics and rubber products NAICS=326, SA	1972:1	2010:8
	Mining			
23		IP:Oil and gas extraction NAICS=211, SA	1972:1	2010:8
24		IP:Mining (except oil and gas) NAICS=212, SA	1972:1	2010:8
25		IP:Support activities for mining NAICS=213, SA	1972:1	2010:8
	Utilities			
26		IP:Electric power generation, transmission and distribution NAICS=2211, SA	1972:1	2010:8
27		IP:Natural gas distribution NAICS=2212, SA	1972:1	2010:8
	Industrial Production by Market			
	Cons Gds: Durables			
28		IP:Automotive products, SA	1959:1	2010:8
29		IP:Autos and trucks, consumer, SA	1967:1	2010:8
30		IP:Auto parts and allied goods, SA	1959:1	2010:8
31		IP:Other durable goods, SA	1959:1	2010:8
32		IP:Computers, video and audio equipment, SA	1967:1	2010:8
33		IP:Appliances, furniture, and carpeting, SA	1967:1	2010:8
34		IP:Miscellaneous durable goods, SA	1959:1	2010:8
	Cons Gds: Nondurables			
35		IP:Foods and tobacco, SA	1959:1	2010:8
36		IP:Clothing, SA	1959:1	2010:8
37		IP:Chemical products, SA	1959:1	2010:8
38		IP:Paper products, SA	1959:1	2010:8
39		IP:Miscellaneous nondurable goods, SA	1972:1	2010:8
40		IP:Consumer energy products, SA	1959:1	2010:8
41		IP:Fuels, SA	1959:1	2010:8
42		IP:Residential utilities, SA	1959:1	2010:8
	Equipment			
43		IP:Transit equipment, SA	1959:1	2010:8
44		IP:Information processing and related equipment, SA	1967:1	2010:8
45		IP:Industrial and other equipment, SA	1967:1	2010:8
46		IP:Industrial equipment, SA	1967:1	2010:8
47		IP:Other equipment, SA	1967:1	2010:8
48		IP:Oil and gas well drilling and manufactured homes, SA	1959:1	2010:8
49		IP:Defense and space equipment, SA	1959:1	2010:8
	Materials: Durables			
50		IP:Consumer parts, SA	1959:1	2010:8
51		IP:Equipment parts, SA	1959:1	2010:8
52		IP:Computer and other board assemblies and parts, SA	1972:1	2010:8
53		IP:Semiconductors, printed circuit boards, and other, SA	1959:1	2010:8
54		IP:Other equipment parts, SA	1967:1	2010:8
55		IP:Other durable materials, SA	1959:1	2010:8
56		IP:Basic metals, SA	1959:1	2010:8

Table A.1 (*continued*)

No.		Series class	Start	End
57		IP:Miscellaneous durable materials, SA	1959:1	2010:8
	Materials: Nondurables			
58		IP:Textile materials, SA	1967:1	2010:8
59		IP:Paper materials, SA	1967:1	2010:8
60		IP:Chemical materials, SA	1967:1	2010:8
61		IP:Other nondurable materials, SA	1959:1	2010:8
62		IP:Containers, SA	1959:1	2010:8
63		IP:Miscellaneous nondurable materials, SA	1967:1	2010:8
	Materials: Energy			
64		IP:Primary energy, SA	1967:1	2010:8
65		IP:Converted fuel, SA	1967:1	2010:8
	NonIndustrial Supplies			
66		IP:Construction supplies, SA	1959:1	2010:8
67		IP:Business supplies, SA	1959:1	2010:8
68		IP:General business supplies, SA	1959:1	2010:8
69		IP:Commercial energy products, SA	1967:1	2010:8
	Employment by industry			
	Mining and logging			
70		Logging	1959:1	2010:9
71		Oil and gas extraction	1972:1	2010:9
72		Mining except oil and gas	1990:1	2010:9
73		Support activities for mining	1990:1	2010:9
	Construction			
74		Construction of Buildings	1990:1	2010:9
75		Heavy and civil engineering construction	1990:1	2010:9
76		Specialty trade contractors	1976:1	2010:9
	Mfg: Durables			
77		Wood products	1990:1	2010:9
78		Nonmetallic mineral products	1959:1	2010:9
79		Primary metals	1990:1	2010:9
80		Fabricated metal products	1990:1	2010:9
81		Machinery	1990:1	2010:9
82		Computer and electronic products	1990:1	2010:9
83		Electrical equipment and appliances	1990:1	2010:9
84		Transportation equipment	1990:1	2010:9
85		Furniture and related products	1990:1	2010:9
86		Miscellaneous manufacturing	1990:1	2010:9
	Mfg: Nondurables			
87		Food manufacturing	1990:1	2010:9
88		Beverages and tobacco products	1990:1	2010:9
89		Textile mills	1990:1	2010:9
90		Textile product mills	1990:1	2010:9
91		Apparel	1990:1	2010:9
92		Leather and allied products	1990:1	2010:9
93		Paper and paper products	1990:1	2010:9
94		Printing and related support activities	1990:1	2010:9
95		Petroleum and coal products	1990:1	2010:9
96		Chemicals	1990:1	2010:9
97		Plastics and rubber products	1990:1	2010:9
	Wholesale trade			
98		Durable goods	1990:1	2010:9
99		Nondurable goods	1990:1	2010:9
100		Electronic markets and agents and brokers	1990:1	2010:9
	Retail trade			
101		Motor vehicle and parts dealers	1990:1	2010:9
102		Furniture and home furnishings stores	1990:1	2010:9
103		Electronics and appliance stores	1990:1	2010:9
104		Building material and garden supply stores	1990:1	2010:9
105		Food and beverage stores	1990:1	2010:9
106		Health and personal care stores	1990:1	2010:9
107		Gasoline stations	1990:1	2010:9
108		Clothing and clothing accessories stores	1990:1	2010:9
109		Sporting goods, hobby, boo, and music stores	1990:1	2010:9
110		General merchandise stores	1990:1	2010:9
111		Miscellaneous store retailers	1990:1	2010:9

(*continued on next page*)

Table A.1 (*continued*)

No.		Series class	Start	End
112		Nonstore retailers	1990:1	2010:9
	Transportation and warehousing			
113		Air transportation	1990:1	2010:9
114		Rail transportation	1959:1	2010:9
115		Water transportation	1990:1	2010:9
116		Truck transportation	1990:1	2010:9
117		Transit and ground passenger transportation	1990:1	2010:9
118		Pipeline transportation	1990:1	2010:9
119		Scenic and sightseeing transportation	1990:1	2010:9
120		Support activities for transportation	1990:1	2010:9
121		Couriers and messengers	1990:1	2010:9
122		Warehousing and storage	1990:1	2010:9
	Utilities			
123		Utilities	1964:1	2010:9
	Information			
124		Publishing industries	1990:1	2010:9
125		Motion picture and sound recording industries	1990:1	2010:9
126		Broadcasting except Internet	1990:1	2010:9
127		Telecommunications	1990:1	2010:9
128		Data processing, hosting and related activities	1990:1	2010:9
129		Other information services	1990:1	2010:9
	Financial activities			
130		Monetary authorities—central bank	1990:1	2010:9
131		Credit intermediation and related activities	1990:1	2010:9
132		Securities, commodities, investments	1990:1	2010:9
133		Insurance carriers and related activities	1990:1	2010:9
134		Funds, trusts, and other financial vehicles	1990:1	2010:9
135		Real estate	1990:1	2010:9
136		Rental and leasing services	1990:1	2010:9
137		Lessors of nonfinancial intangible assets	1990:1	2010:9
	Professional and business services			
138		Professional and technical services	1990:1	2010:9
139		Management of companies and enterprises	1990:1	2010:9
140		Administrative and waste services	1990:1	2010:9
	Education and health services			
141		Education services	1990:1	2010:9
142		Health care	1990:1	2010:9
143		Social assistance	1990:1	2010:9
	Leisure and hospitality			
144		Arts/entertainment/recreation	1990:1	2010:9
145		Accommodation	1972:1	2010:9
146		Food services and drinking places	1990:1	2010:9
147		Other services	1959:1	2010:9
	Government			
148		Federal	1959:1	2010:9
149		State	1959:1	2010:9
150		Local	1959:1	2010:9
	Employment at higher level of aggregation			
	Manufacturing			
151		Durables	1959:1	2010:9
152		Nondurables	1959:1	2010:9
153		Construction	1959:1	2010:9
	Services			
154		Education and health	1959:1	2010:9
155		Financial activities	1959:1	2010:9
156		Government	1959:1	2010:9
	Services			
157		Information	1959:1	2010:9
158		Leisure and hospitality	1959:1	2010:9
159		Professional and bus services	1959:1	2010:9
160		Other services	1959:1	2010:9

Table A.1 (*continued*)

No.		Series class	Start	End
	Nat. Resources and Mining			
161		Nat. resources and mining	1959:1	2010:9
	Trade			
162		Retail	1959:1	2010:9
163		Wholesale	1959:1	2010:9
164		Trans/Utilities (USTPU-USTRADE-USWTRADE)	1959:1	2010:9
	Manufacturing and trade sales (SIC classification)			
	Mfg: Durables			
165		Lumber and wood products	1967:1	1996:12
166		Furniture and fixtures	1967:1	1996:12
167		Stone, clay, and glass products	1967:1	1996:12
168		Primary metals	1967:1	1996:12
169		Fabricated metals	1967:1	1996:12
170		Industrial machinery	1967:1	1996:12
171		Electronic machinery	1967:1	1996:12
172		Transportation equipment	1967:1	1996:12
173		Instruments	1967:1	1996:12
174		Other manufacturing	1967:1	1996:12
	Mfg: Nondurables			
175		Food and kindred products	1967:1	1996:12
176		Tobacco products	1967:1	1996:12
177		Textile mill products	1967:1	1996:12
178		Apparel products	1967:1	1996:12
179		Paper and allied products	1967:1	1996:12
180		Printing and publishing	1967:1	1996:12
181		Chemical and allied products	1967:1	1996:12
182		Petroleum products	1967:1	1996:12
183		Rubber and plastic products	1967:1	1996:12
184		Leather and leather products	1967:1	1996:12
	Merchant wholesale: durable goods			
185		Motor vehicles	1967:1	1996:12
186		Furniture and furnishings	1967:1	1996:12
187		Lumber and construction	1967:1	1996:12
188		Professional and commercial	1967:1	1996:12
189		Metals and minerals	1967:1	1996:12
190		Electrical goods	1967:1	1996:12
191		Hardware and plumbing	1967:1	1996:12
192		Machinery, equipment, and supplies	1967:1	1996:12
193		Other durable goods	1967:1	1996:12
	Merchant wholesale: nondurable goods			
194		Paper products	1967:1	1996:12
195		Drugs and sundries	1967:1	1996:12
196		Apparel and piece goods	1967:1	1996:12
197		Groceries	1967:1	1996:12
198		Farm products	1967:1	1996:12
199		Chemical and allied products	1967:1	1996:12
200		Petroleum products	1967:1	1996:12
201		Alcoholic beverages	1967:1	1996:12
202		Other nondurable goods	1967:1	1996:12
	Retail trade: durable goods			
203		Automotives	1967:1	1996:12
204		Lumber and building stores	1967:1	1996:12
205		Furniture and furnishings	1967:1	1996:12
206		Other durable goods	1967:1	1996:12
	Retail trade: nondurable goods			
207		Food stores	1967:1	1996:12
208		Apparel stores	1967:1	1996:12
209		Department stores	1967:1	1996:12
210		Other general merchandise stores	1967:1	1996:12
	Manufacturing and trade sales (NAICS classification)			
	Mfg: Durables			
211		Wood product manufacturing	1997:1	2010:7
212		Nonmetallic mineral product manufacturing	1997:1	2010:7
213		Primary metal manufacturing	1997:1	2010:7
214		Fabricated metal product manufacturing	1997:1	2010:7
215		Machinery manufacturing	1997:1	2010:7

(*continued on next page*)

Table A.1 (*continued*)

No.		Series class	Start	End
216		Computer and electronic product manufacturing	1997:1	2010:7
217		Electrical equipment, appliance, and component manufacturing	1997:1	2010:7
218		Transportation equipment manufacturing	1997:1	2010:7
210		Furniture and related product manufacturing	1997:1	2010:7
220		Miscellaneous durable goods manufacturing	1997:1	2010:7
	Mfg: Nondurables			
221		Food manufacturing	1997:1	2010:7
222		Beverage and tobacco product manufacturing	1997:1	2010:7
223		Textile mills	1997:1	2010:7
224		Textile product mills	1997:1	2010:7
225		Apparel manufacturing	1997:1	2010:7
226		Leather and allied product manufacturing	1997:1	2010:7
227		Paper manufacturing	1997:1	2010:7
228		Printing and related support activities	1997:1	2010:7
229		Petroleum and coal product manufacturing	1997:1	2010:7
230		Chemical manufacturing	1997:1	2010:7
231		Plastics and rubber product manufacturing	1997:1	2010:7
	Merchant wholesale industries: durable goods			
232		Motor vehicles, parts, and supplies wholesalers	1997:1	2010:7
233		Furniture and home furnishings wholesalers	1997:1	2010:7
234		Lumber and other construction materials wholesalers	1997:1	2010:7
235		Professional and commercial equipment wholesalers	1997:1	2010:7
236		Metal and mineral (except petroleum) wholesalers	1997:1	2010:7
237		Electrical goods wholesalers	1997:1	2010:7
238		Hardware and plumbing and heating equipment wholesalers	1997:1	2010:7
239		Machinery, equipment, and supplies wholesalers	1997:1	2010:7
240		Miscellaneous durable goods wholesalers	1997:1	2010:7
	Merchant wholesale industries: nondurable goods			
241		Paper and paper products wholesalers	1997:1	2010:7
242		Drugs and druggists' sundries wholesalers	1997:1	2010:7
243		Apparel, piece goods, and notions wholesalers	1997:1	2010:7
244		Grocery and related products wholesalers	1997:1	2010:7
245		Farm product raw material wholesalers	1997:1	2010:7
246		Chemical and allied products wholesalers	1997:1	2010:7
247		Petroleum and petroleum products wholesalers	1997:1	2010:7
248		Beer, wine, and distilled alcoholic beverages wholesalers	1997:1	2010:7
249		Miscellaneous nondurable goods wholesalers	1997:1	2010:7
	Retail trade industries			
250		Motor vehicle and parts dealers	1997:1	2010:7
251		Furniture, furnishings, electronics, and appliance stores	1997:1	2010:7
252		Building material and garden equipment and supplies dealers	1997:1	2010:7
253		Food and beverage stores	1997:1	2010:7
254		Clothing and clothing accessories stores	1997:1	2010:7
255		General merchandise stores	1997:1	2010:7
256		Other retail stores	1997:1	2010:7
	Personal income (all series are deflated by the PCE deflator)			
	Wages and salaries			
257		Manufacturing (SIC)	1959:1	2000:12
258		Distributive industries (SIC)	1959:1	2000:12
259		Service industries (SIC)	1959:1	2000:12
260		Manufacturing (NAICS)	2001:1	2010:8
261		Trade, transportation, and utilities (NAICS)	2001:1	2010:8
262		Other services—producing industries (NAICS)	2001:1	2010:8
263		Government	1959:1	2010:8
264		Supplements to wages and salaries	1959:1	2010:8
	Prop. income			
265		Farm	1959:1	2010:8
266		Nonfarm	1959:1	2010:8
	Rental income			
267		Rental income	1959:1	2010:8
	Personal income receipts on assets			
268		Interest	1959:1	2010:8
269		Dividend	1959:1	2010:8
270		Personal current taxes	1959:1	2010:8

essentially flat from 2000:9–2001:3, with a slight local peak in 2000:9, and monthly GDP(Avg) is increasing with only minor fluctuations over this period and into the summer of 2001. Two interpretations of the Burns–Mitchell method seem possible in this circumstance. The first would be to select the end of this long flat episode, which would accord with the NBER date of 2001:3. This estimate is later than any of those in Table 3 because Table 3 does not consider end-of-episode dating. However, an end-of-episode dating rule would lag the NBER dates: the end-of-episode dates in the three cases considered here postdate the weighted mode, which on average lags the NBER chronology by only 0.13 months. The second interpretation, which is the approach we have adopted in this paper, would choose the mode of the second cluster, which (with a large spike in the histogram) is 2000:12.

The 2007:12 peak. Aside from monthly GDI, the average-then-date chronologies all estimate the 2007:12 peak to be within a month of the NBER date. In contrast, the weighted mode estimator places the 2007:12 peak six months earlier than the NBER date, with a tight standard error (1.1). Inspection of the histogram and weighted kernel density estimate for this episode, shown in Fig. 8, reveals however that this episode has an interesting pattern of turning points of the disaggregated series. The episode has a long period (approximately 12 months) over which many series turn, and the kernel density estimator has two modes, the higher one being 6 months before the NBER date and a slightly lower one two months after. Thus inspection of the histogram and weighted density estimate suggests that in this episode Burns and Mitchell's (1946) "close of transition period" rule might apply, which would place the turning point a month or two after the NBER peak. We do not have a mathematical implementation for the close-of-transition-period concept so we cannot provide a more precise estimate for the 2007:12 peak, or a standard error, based on that approach, however that concept does suggest placing substantially less weight on the weighted mode estimator at this turning points than at the other turning points.

5. Discussion

The empirical results in Section 4 suggest that the date-then-average procedures, including the new confidence intervals for reference cycle turning points, have the potential to provide useful information to supplement the process of determining reference cycle chronologies. The modal turning point is closely related to the approach used by Burns and Mitchell (1946) and the early NBER researchers. In addition to this historical link, the use of the mode seems to be preferable empirically to the mean because the mean is sensitive to outliers as was noted in the discussion of the 1991:3 trough.

The exercise here focuses on subaggregates within only four classes of series. Arguably more classes should be considered, indeed the early NBER researchers considered a much broader set of series than these four. For example, the importance of GDP as a measure of output suggests extending this analysis to include monthly subaggregates of GDP when available.

While we have been able to produce standard errors for the date-then-average chronologies, no such standard errors exist for the average-then-date chronologies. Developing a frequentist distribution theory for Bry–Boschan turning points of a single aggregate time series remains an intriguing research problem.

Acknowledgments

This research was funded in part by NSF grant SBR-0617811. We thank Marcelle Chauvet, Massimilliano Marcellino, and a referee for helpful comments, and Carolin Pflueger and Vania Stravrakeva for research assistance.

Appendix A. The disaggregated data set

See Table A.1.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jeconom.2013.08.034.

References

Bry, Gerhard, Boschan, Charlotte, 1971. Cyclical Analysis of Time Series: Procedures and Computer Programs. NBER, New York.

Burns, Arthur F., Mitchell, Wesley C., 1946. Measuring Business Cycles. NBER, New York.

Chauvet, Marcelle, Piger, Jeremy, 2008. A comparison of the real-time performance of business cycle dating methods. Journal of Business and Economic Statistics 26, 42–49.

Hamilton, James D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357–384.

Hamilton, James D., 2011. Calling recessions in real time. International Journal of Forecasting 27 (4), 1006–1026.

Harding, Donald, Pagan, Adrian, 2006. Synchronization of cycles. Journal of Econometrics 132, 59–79.

Nalewaik, Jeremy J., 2010. The income- and expenditure-side estimates of U.S. output growth. In: Brooking Papers on Economic Activity, 2010. Spring, pp. 71–106.

NBER Business Cycle Dating Committee, 2008. Determination of the December 2007 Peak in Economic Activity, at http://www.nber.org/cycles/dec2008.html.

Parzen, Emmanuel, 1962. On estimation of a probability density function and mode. Annals of Mathematical Statistics 33, 1065–1076.

Romano, J.P., 1988. On weak convergence and optimal kernel density estimates of the mode. Annals of Statistics 16, 629–647.

Stock, James H., Watson, Mark W., 1989. New indexes of coincident and leading economic indicators. In: Macroeconomics Annual, Vol. 4, 1989. M.I.T. Press.

Stock, James H., Watson, Mark W., 2010a. New Indexes of Monthly GDP, available at http://www.princeton.edu/~mwatson/mgdp_gdi.html.

Stock, James H., Watson, Mark W., 2010b. Indicators for dating business cycles: cross-history selection and comparisons. American Economic Review: Papers and Proceedings 16–19.

Ziegler, Klaus, 2003. On the asymptotic normality of kernel regression estimators of the mode in the nonparametric random design model. Journal of Statistical Planning and Inference 115, 123–144.