business cycle measurement

Don Harding and Adrian Pagan
Edited by Steven N. Durlauf and Lawrence E. Blume

Abstract

We describe different ways of measuring the business cycle. Institutions such as the NBER, OECD and IMF do this by locating the turning points in series taken to represent the aggregate level of economic activity. The turning points are determined according to rules that either come from a parametric model or are nonparametric. Once located, information can be extracted on cycle characteristics. We also distinguish between cases where single and multiple series are used to represent the level of activity.

Keywords

Burns, A.; business cycle; business cycle measurement; censoring operations; coincident indices; crossing points; data filters; fluctuations vs cycles; growth cycles; Markov switching (MS) processes; Mitchell, W.; periodic cycles; random variables; reference cycle; spectral analysis; turning points

Article

Measurement of business cycles provides a reference point against which macroeconomic theories and policy discussion can be assessed. The process requires an operational definition of a cycle, criteria to distinguish business cycles from other forms of fluctuation, procedures to detect the presence of a business cycle, and methods to measure its features. A central theme of this entry is that good measurement should not prejudge the nature of the phenomena under investigation. Moreover, it should produce statistics which are informative about features of interest and which can be formally analysed.

Defining and detecting cycles

In their classic work Measuring Business Cycles, Burns and Mitchell (BM) (1946) define specific cycles in a series $y_t$ in terms of turning points in its sample path. This tradition has been central to work at the NBER and other institutions such as the IMF (2002) and the OECD (leading indicators). When it came to discussing the business cycle, BM simply referred to $y_t$ as the level of aggregate economic activity, although in this article we will regard it as the log of economic activity, as the turning points in the level and the log of economic activity are the same. When Mintz (1969; 1972) had trouble finding turning points in the level of activity in surging economies such as West Germany’s, this led her to first extract a permanent component $p_t$ from $y_t$ and to then study turning points in $z_t = y_t - p_t$. The resulting growth cycle in $z_t$ has many forms depending on the method used to extract the permanent component. Others, such as the Economic Cycle Research Institute (ECRI) (growth rate cycle), have studied turning points in the differenced data $\Delta y_t$. A generalization of this, explored by Kedem (1980; 1994) and Harding (2003), is to study turning points in $\Delta' y_t$.

At the time Mitchell began his work, the alternative way of thinking about cycles (or oscillations) was to view $y_t$ as composed of periodic components represented by sine and cosine waves, that is

$$y_t = \sum_{j=1}^{\infty} \alpha_j \cos(\lambda_j t) + \beta_j \sin(\lambda_j t),$$

(1)

where $\lambda_j$ is the frequency of the $j^\text{th}$ oscillation. If $\tau_n = 1$ there would be a single periodic cycle. The problem with this way of looking at
cycles was that few economic time series showed evidence of periodicity. To overcome that problem \( \alpha_j \) and \( \beta_j \) were allowed to vary stochastically over time. Specifically, they were treated as uncorrelated random variables with zero mean and variance \( \sigma_j^2 \). This formulation meant that \( y_t \) had to be a stationary random variable and so could not be applied to the levels of variables such as GDP (unlike turning point analysis). However, in this form one can measure the importance of the \( j^{th} \) periodic cycle by looking at the ratio of \( \sigma_j^2 \) to the variance of \( y_t \) and it is the basis of spectral analysis. Such a perspective has increasingly been referred to as studying fluctuations rather than cycles, since the focus of attention is upon the variance of \( y_t \).

To understand the difference between these alternative ways of measuring cycles, take the special case where \( \lambda_1 = 0 \) and there is another frequency \( \lambda_2 \). Then

\[
y_t = \alpha_2 \cos \lambda_2 t + \beta_2 \sin \lambda_2 t + \epsilon_t = y_t' + \epsilon_t.
\]

(2)

Now there are certainly turning points in the series \( y_t' \) and the period between them is determined by \( \lambda_2 \). In contrast, the turning points in \( y_t \) will also be affected by the random variable \( \epsilon_t \), and this may be very different to those in \( y_t' \). Information about cycles gathered from spectral analysis concerns the nature of turning points in \( y_t' \) and not \( y_t \). To give a more concrete illustration of this point, suppose that the model for \( y_t \) is of the form

\[
y_t = 1.4 y_{t-1} - .53 y_{t-2} + \epsilon_t.
\]

Then the periodic cycle in \( y_t \) can be isolated by setting \( \epsilon_t = 0 \) to get \( y_t' \). To use the dating methods of an institution like NBER, the turning points in \( y_t' \) are 22 quarters apart, as could also be discovered by computing the roots of \( (1 - 1.4L + .53L^2) = 0 \). However, applying the same methods to \( y_t \), one finds that the turning points in \( y_t \) will be on average 12 quarters apart. A further disadvantage of the periodic cycle approach is that the data needs to be filtered to render it stationary before analysis proceeds and, as Cogley observes elsewhere in this dictionary (data filters), the filters most commonly used by macroeconomists can introduce spurious periodic cycles, thereby blurring the picture.

**Locating turning points**

To locate turning points in a series it is necessary to define what these are and to provide some way of recognizing them in a given data-set. An obvious solution is to use the idea that peaks (troughs) are local maxima (minima) in the series \( y_t \). Hence, if \( \forall_t (A_t) \) are binary variables taking the value of unity where there is a peak (trough) at \( t \) and zero otherwise, applying the proposed definition gives

\[
\forall_t = 1(y_t < y_{t \pm j}, 1 \leq j \leq k)
\]

(3)

\[
A_t = 1(y_t > y_{t \pm j}, 1 \leq j \leq k),
\]

(4)

In eqs. (3) and (4) \( 1(A) \) is the indicator function taking the value 1 if the event \( A \) is true and zero otherwise. Of course, this still leaves one with the need to describe the interval over which the local maxima or minima are said to occur, that is, a choice needs to be made regarding
k. To replicate the main features of Burns and Mitchell's specific cycle dating procedures, it is necessary to set $k \leq 5$ for monthly data or $k = 2$ for quarterly data.

This is not the last of the choices that need to be made when locating turning points, but the others do not relate to the location of local maxima and minima. Rather, they concern the question of whether one should eliminate some of the local turns in deciding on a final set of turning points. Mostly these extra restrictions are imposed as phase length constraints, where phases are the periods of expansions and contractions between turning points. Thus, NBER dating procedures require that completed phase and complete cycles durations last longer than 5 and 15 months respectively. These are generally referred to as censoring operations. Whether turning points should be censored depends on the objectives of the research. If the objective is to match NBER business cycle dates, then censoring is essential. But if the researcher is pursuing other objectives such censoring may not be necessary. Censoring turning points makes it much harder to formally analyse the statistics produced and this may provide an important reason for not imposing them.

BM acknowledged that the final set of dates they selected for turning points reflected considerable amounts of judgement and incorporated specific information about economic activity at particular dates. Today, academic economists are primarily interested in the average characteristics of the cycle, and so it may well be that automated methods of turning point detection become attractive. In the early post-Second World War period many of the procedures used by BM were codified, producing an expert system for locating turning points. Ultimately, Bry and Boschan (1971) produced an algorithm and FORTRAN program (called BB here) that largely replicated this expert system. Subsequently Mark Watson (1994) implemented this algorithm in the language GAUSS, and that code is available at (http://www.princeton.edu/~mwatson).

There were three key components to the BB algorithm. The first was to engage in some smoothing of the series and to find an initial set of turning points using eqs. (3) and (4) with $k = 5$. The second was to eliminate enough of these turning points so as to ensure that expansion and contraction phases exceeded 5 months in duration, while completed cycles exceed 15 months in duration. The third component was to ensure that peaks and troughs are deleted by deleting multiple sequential occurrences of these. That was done through the application of various rules, such as choosing between two peaks based on which had the higher value of $y_j$.

Although BB were interested in analysing monthly data, they suggested a method for working with quarterly data that involved treating the observations on each of the months in a quarter as one-third of the quarterly value. A variant of BB has been developed by Harding and Pagan (2002) and called BBQ. It omits the smoothing in the BB algorithm but retains the three key principles of the BB algorithm. It also sets $k = 2$ and makes the minimum phase and cycle lengths two and five quarters respectively. Faster recursive algorithms for locating turning points have been developed by Arts, Marcellino and Proietti (2004) and James Engel. Engel's computer programs are called MBBQ. They are written in MATLAB and GAUSS and are available at the National Centre for Econometric Research (MBBQ Code).

Model-based procedures for defining and locating turning points

The procedures above do not require any knowledge of the data-generating process for $y_j$. An alternative approach is to adopt a model of $\Delta y_j$ and use this to locate turning points. To date the models used are parametric and generally feature two regimes. Perhaps the best known parametric model is that of Hamilton (1989), where the growth rate is treated as a Markov switching (MS) process of the form $\Delta y_j = \mu_0(1 - 2z_t) + \mu_zz_t + e_t$. Here $\mu_0$ are the growth rates in the two regimes, and these are indexed by a latent binary state, $z_t$, while $e_t$ is a normally distributed zero mean error term. Here $\mu_0$ is the growth rate of the low growth state and $\mu_1$ is the high growth rate. Sometimes the restriction $\mu_0 < 0$ is also imposed. The model is completed by specifying the transition probabilities of moving from $z_{t-1} = 0$ or 1 to $z_t = 1$ or 0. The model can be made more complex with extra dynamics, different variances in each regime, allowing the transition probabilities to depend on some observable data, and so on. This parametric model is used to compute the conditional probability, $\Pr(\xi \mid A_t)$, where $A_t$ is either all or a subset of the growth rates $\{\Delta y_j\}_{j=1}^T$. Thus the estimate of $\Pr(\xi = 1 \mid A_t)$ is a function of whatever growth rates are in $A_t$. Generally this probability will be a nonlinear function of the elements in $A_t$, although a linear function can be quite a good approximation—see Harding and Pagan (2003) for an example.

The cycle is then associated with a binary variable $S_t$ that takes the value 1 in expansion and zero in contraction. A rule is used to construct $S_t$ by comparing the estimated probability of being in the high growth state with some critical value. Hamilton chose .5, and most of those using the technique have followed suit. Consequently, if $\Pr(\xi = 1 \mid A_t) > .5$, an expansion is signified and $S_t$ is set to unity. If the criterion is not satisfied $S_t$ is set to zero. Notice that the $z_t$ are not the phase states; the latter are $S_t$. They are simply a device for producing some nonlinear structure in $\Delta y_j$, although often one can think of the outcomes for $z_t$ as signifying a low or high growth period. The correlation between $S_t$ and $\xi_t$ may be very low. Many applications of this methodology have now been made and the MS model that one chooses seems to vary a lot with the series it is being applied to. The simple one described above rarely works satisfactorily.
In most instances a decision about the utility of the method is made by comparing the business cycle states produced by the rule based on the magnitude of $P_{Y^i_t = 1 | A_t} > .5$ with those found by turning point methods. Because of the latter comparison one has to ask what the advantages there are in using a model to locate turning points. Chauvet and Piger (2003) claim that an advantage of the model-based approach is that it allows an investigator to forecast turning points in real time. There is some truth to this but it is exaggerated. Since forecasts can be found for any such model, they could be passed through any chosen dating algorithm to determine the predicted phases.

**Measuring cycle features**

Turning points segment time series into phases. An expansion phase runs from the trough to the next peak. A contraction runs from a peak to the next trough. In what follows it is easiest to just describe the derivation of information on expansions.

The two most basic statistics related to phases are duration and amplitude. The *duration* of an expansion is the number of periods of time between the trough and next peak. The *amplitude* of an expansion measures the change in $y_t$ from trough to the next peak. In many cases $y_t$ is the log of some variable such as GDP or industrial production, that is, $y_t = \ln(Y_t)$, and the amplitude has a natural interpretation as the approximate percentage change in $Y_t$ between trough and peak.

Duration and amplitude form two sides of a triangle. Connecting the trough and peak produces the hypotenuse. If $y_t = h(\gamma_t)$, then the hypotenuse represents the path followed by a variable that exhibits a constant growth rate during an expansion. With this in mind it is instructive to inspect the actual path followed by the data, and to compare that path with the constant growth path represented by the hypotenuse. Figure 1 shows how US expansion paths have deviated from the constant growth rate path in the post-Second World War period. The important feature evident in this figure is that the growth rate of GDP is not constant over the expansion phase and typically is highest in the first half of an expansion.

Figure 1

While comparisons such as that in Figure 1 are visually informative, there is also a need for statistics that summarize the average shape of
phases. Sichel (1994) divides expansions into three stages, computes the average growth rate for each stage, and shows graphs of these, as well as providing formal statistical tests of equality of the growth rates in each stage. Harding and Pagan (2002) compare the cumulated gain in an expansion with what it would have been if growth had been constant throughout the phase. This comparison was motivated by the idea mentioned above, that a plot of $y_t$ against $t$ during an expansion would look like a triangle if growth had been constant. The area of such a triangle would be one-half the product of the amplitude and duration. If growth was not constant the area under the path actually followed by activity during the expansion would differ from the triangle. Thus, a comparison of the two areas provides a measure of the extent of departure from a constant growth scenario. The evidence seems to be that expansions do not feature constant growth in some countries like Australia, the United States and the UK, but do so in many Continental European countries. The shape analysis is interesting since a linear process for $\Delta y_t$ will produce phases that, on average, have constant growth rates. So a failure to see this signals the need for a nonlinear process for $\Delta y_t$. The shape analysis also provides a useful tool for testing whether nonlinear models produce realistic business cycles.

All of the methods for summarizing business cycle information can be applied to growth cycles and to data that have undergone higher-order differencing. In addition, Sichel (1993) suggested tests for ‘deepness’ and ‘steepness’ in the growth cycle that were effectively tests for symmetry in the densities of $x_t$ and $\Delta x_t$.

Using multivariate information in defining and detecting business cycles

Burns and Mitchell's famous definition of a business cycle – ‘Business cycles are a type of fluctuation found in the aggregate economic activity of nations...a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general...contractions...' (1946, p. 1) – has two aspects. One points to the need to identify aggregate economic activity, and the other to the fact that there should be synchronization across many series during the phases of a business cycle. Burns and Mitchell commented that GDP was a suitable index of economic activity, although others, such as Moore and Zarnowitz (1986), have preferred a weighted average of several series rather than a single one. However, since data on GDP was not available to Burns and Mitchell, for either the time period or the frequency in which they were interested, it is natural that they placed more emphasis upon the second component of their definition when discussing the business cycle.

This second component emphasizes synchronization of the cycles in the specific series taken to represent economic activity. Burns and Mitchell took the turning points in many series and then extracted a reference cycle by determining those dates which peaks and troughs ‘clustered around’. So a primary task is to be able to measure the tightness of the clusters. At the end of the process one also wishes to know how synchronized each of the specific cycles is with the cycle in the aggregate.

Harding and Pagan (2006) develop procedures to measure the tightness of clusters of turning points and the degree of synchronization of cycles through concordance indices that measure the fraction of time spent in the same phase. They apply those procedures to the series referred to by the NBER when dating the business cycle, and find that the turning points in those series are tightly clustered together. Harding (2003) finds that between March 1949 and September 2001 there is a concordance of 0.96 between the NBER business cycle states and the cycle obtained by locating turning points in US GDP.

Automated construction of the reference cycle

To automate the calculation of the reference cycle requires some rules which will distill the specific cycle turning points into a single set of turning points. To determine what these rules might be, one could look at the NBER Business Cycle Dating Committee procedure. It has a similar modus operandi to that of Burns and Mitchell, as seen in its discussion about dating the 2001 recession (NBER, 2003). However, one rarely gets a precise description either of how its decisions are made or of the series used in that process. In addition, it seems as if the series which have been most influential in decisions may have been different at different periods in time. The clearest description of the procedures for aggregating turning points in a set of series to create a reference cycle is in Boehm and Moore (1984), who explain how NBER methods were used when establishing a reference cycle for Australia. Their description can be taken as authoritative because Moore was a pivotal figure in the NBER Business Cycle Dating Committee for many years. Moore and Zarnowitz (1986) also provide information on methods used by NBER in dating the business cycle.

Given that the process for establishing the reference cycle is a little vague, it should not be surprising that there have been few attempts at producing automated dating algorithms to establish it from multivariate series. Harding and Pagan (2006) construct an algorithm to replicate the NBER procedures described by Boehm and Moore (1984). They obtain the ‘clustering parameter’ which is essential to measuring the tightness of turning point clusters by looking at Boehm and Moore's spreadsheets. The resulting algorithm has produced a reference cycle that matches the Australian version established by Boehm and Moore quite well. Subsequently, it has been tested on US data, and is able to produce quite a good replication of the reference cycle for the United States, even though the clustering parameter had been calibrated with Australian data.
Model-based procedures for defining detecting and extracting a reference cycle

Recently, academic economists have used parametric models to construct a coincident index and the reference cycle from $n$ multivariate series $\Delta y_{1n}, \ldots, \Delta y_{nT}$. A common element to all approaches is to write $\Delta y_{jt}$ as a function of a common component $f_t$ and idiosyncratic components $u_{jt}, (j = 1, \ldots, n)$. Hence a simple representation would be $\Delta y_{jt} = f_t + u_{jt}$. The $f_t$ is often thought of as the coincident index of the business cycle. Of course, there may be more than one $f_t$ but, ultimately, we can think of combining them to form a single variable. There are then many ways that models for $f_t$ and $u_{jt}$ might be specified, depending upon how strong the assumptions are that one wishes to make about the nature of $f_t$ and $u_{jt}$. Often $f_t$ is given an MS form (for example, Chauvet and Piger, 2003). Depending on what these assumptions are, they will determine how an estimate of $f_t$ is to be made. Stock and Watson (1991) and Chauvet (1998) represent different approaches. In some instances one can avoid specifying precise parametric models for $f_t$ and $u_{jt}$, restricting them only to be in a general class. Forni et al. (2001)'s dynamic factor approach is the main representative of this latter technique. The main issue with these approaches is that the coincident index and reference cycle obtained are conditioned on the assumptions made about the data-generating process. For that reason these approaches cannot provide a neutral measurement of the reference cycle.

Conclusion

Although widely used in official circles, Burns and Mitchell's methods of measuring cycles through turning points have been less popular in academia. But this has changed in recent years. There are a number of reasons why the methods have become increasingly attractive. First, information about the nature of the cycle phases can be generated, and this shape information proves important when one tries to construct models of economic activity. Second, the literature now contains expert systems for locating turning points, and these have been coded into various computer languages, thereby eliminating the judgmental aspect of the method. Nevertheless, the automatically generated turning points have been quite good approximations to those found via judgment. Third, the ability to produce simulated data from parametric models means that such information can be passed through the algorithms for locating turning points to produce simulated distributions for the statistics that summarize the features of the cycle. Fourth, the emerging mathematics literature on crossing points provides a natural foundation on which to build a distribution theory for Burns and Mitchell's methods. Fifth, there is now a large literature on parametric methods for locating turning points and measuring cycles. This latter literature can readily be linked to the nonparametric turning point approach of investigators such as Burns and Mitchell, as seen in Harding and Pagan (2003).

See Also

- Burns, Arthur Frank
- data filters
- international real business cycles
- Mitchell, Wesley Clair

Bibliography


**How to cite this article**