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

[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).

* 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

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.