A new approach to dating the reference cycle *

Maximo Camacho† University of Murcia

María Dolores Gadea‡ University of Zaragoza

Ana Gómez Loscos§ Bank of Spain

May 21, 2020

Abstract

This paper proposes a new approach to the analysis of the reference cycle turning points, defined on the basis of the specific turning points of a broad set of coincident economic indicators. Each individual pair of specific peaks and troughs from these indicators is viewed as a realization of a mixture of an unspecified number of separate bivariate Gaussian distributions whose different means are the reference turning points. These dates break the sample into separate reference cycle phases, whose shifts are modeled by a hidden Markov chain. The transition probability matrix is constrained so that the specification is equivalent to a multiple change-point model. Bayesian estimation of finite Markov mixture modeling techniques is suggested to estimate the model. Several Monte Carlo experiments are used to show the accuracy of the model to date reference cycles that suffer from short phases, uncertain turning points, small samples and asymmetric cycles. In the empirical section, we show the high performance of our approach to identifying the US reference cycle, with little difference from the timing of the turning point dates established by the NBER. In a pseudo real-time analysis, we also show the good performance of this methodology in terms of accuracy and speed of detection of turning point dates.

Keywords: Business cycles, Turning points, Finite mixture models.
JEL Classification: E32, C22, E27.

*We thank the editor and two referees for helpful comments and suggestions. M. Camacho and M.D. Gadea are grateful for the support of grants PID2019-107192GB-I00 (AEI/10.13039/501100011033) and 19884/GERM/15, and ECO2017-83255-C3-1-P and ECO2017-83255-C3-3-P (MICINU, AEI/ERDF, EU), respectively. We thank Gary Koop for sharing his GAUSS code to determine the number of breaks. All remaining errors are our responsibility. Data and codes that replicate our results are available from the authors' websites. The views in this paper are those of the authors and do not represent the views of the Bank of Spain or the Eurosystem.

†Department of Quantitative Analysis, University of Murcia. Campus de Espinardo 30100 Murcia (Spain). Tel: +34 868 887982, fax: +34 868 887905 and e-mail: mcamacho@um.es

‡Department of Applied Economics, University of Zaragoza. Gran Vía, 4, 50005 Zaragoza (Spain). Tel: +34 9767 61842, fax: +34 976 761840 and e-mail: lgadea@unizar.es

§Bank of Spain, Alcalá, 48, 28014 Madrid (Spain). Tel: +34 91 3385817, fax: +34 915310059 and e-mail: agomezloscos@bde.es

1 Introduction

Since the early work of Burns and Mitchell (1946), a classical problem in business cycle analysis has been to make inferences about the reference cycle dates. These authors postulate that reference cycle turning point dates can be identified by searching for clusters in specific cycle turning point dates from a set of disaggregated coincident economic indicators, which typically cluster around periods of recoveries and declines. In practice, although the distribution of the turning points of the disaggregated series are routinely achieved by employing dating algorithms such as the monthly Bry and Boschan (1971) and the quarterly Harding and Pagan (2002) procedures, the main challenge of this date-then-average literature is to determine the aggregated reference cycle based on these specific dates.[1]

From a set of series that are believed to be roughly coincident, Burns and Mitchell (1946) seek to identify clusters of specific turning points by visual inspection. The authors mark off the zone within which a succession of specific turning points cluster together with a central tendency and proceed to refine the approximate dates based on expert judgment by a sequential process of trial and error. Therefore, the procedure depends on judgments that inevitably condition the reference dates and make it difficult to replicate and update.

Stock and Watson (2010a) provide an interesting contribution in this literature by computing the reference cycles as the means of individual series of turning points. Assuming a segmentation of the time span into business cycle episodes, the data have a standard panel data structure. Then, the parameters can be estimated by ordinary least squares in a specification with fixed effects, an unbalanced panel and missing observations. Although the method contributes by producing standard errors for the estimated reference cycle turning point dates, it takes the sequence of business cycles as given, which considerably limits its empirical implementation.

Stock and Watson (2014) innovate by considering turning points as population concepts. In their approach, turning points are local measures of the central tendency of the population distribution of the set of disaggregated coincident economic indicators. In particular, they focus on the mode as the estimator of the tendency of the kernel density of the sample of individual turning point dates, conditional on the occurrence of a single phase shift in a given episode covering a known time interval.[2] This framework provides the asymptotic distributions of the estimator, which allows them to make statistical inference about the estimated reference cycle. However, they acknowledge that a natural extension of their method would be to examine estimators that simultaneously determine whether there is a turning point and, if so, the date of the turning point.

Harding and Pagan (2016) propose a modification of the algorithm developed by Harding and Pagan (2006), which consists of a set of specific rules and censoring procedures that find the dates that minimize a measure of the average distance between the candidates of reference cycle dates and the turning points in individual series. Contrary to the Stock and Watson's proposals, the Harding and Pagan's algorithm does not need the idea of a pre-specified episode in the reference cycle. Chauvet and Piger (2008) examine the ability of this algorithm to identify the NBER-referenced

[1] On the contrary, the average-then-date approach focuses on one (or a few) highly aggregated time series to date the reference cycle.

[2] They define the different episodes as the NBER turning point dates ± 12 months.

business cycle turning point dates in real time.

This paper addresses the issue of determining the reference cycle by proposing an alternative approach that, like the Harding and Pagan algorithm, overcomes the drawbacks of the existing methods. Our approach assumes that each individual pair of peak and trough dates is a realization of a mixture of an unspecified number of separate bivariate Gaussian distributions. The reference cycle dates are viewed as the means of these distributions, around which the specific dates are clustered, breaking the time span of interest into segments corresponding to the periods occupied by the business cycles. The phase shifts across these cycles are modeled by a discrete-time, discrete-state Markov process with the non-ergodic transition probability matrix constrained so that the model is equivalent to a multiple change-point model following the lines suggested by Chib (1998).

Therefore, our research contributes to the literature in several important ways. First, the aggregate turning points are viewed as population concepts as in Stock and Watson (2010, 2014), in our case, the distinct means of the separate Gaussian distributions. However, as in Harding and Pagan's algorithm, the estimates are not conditional on the known occurrence of a phase shift. Our proposal estimates through the data both the number and the dates of the turning points in a single step. Second, the method is also able to make statistical inference about the reference cycle turning points. Consequently, one can evaluate the uncertainty in the determination of the turning point dates by, for example, computing their confidence or credible intervals. Third, we extend the univariate multiple change-point model of Chib (1998) to a multivariate framework.

We evaluate the ability of the mixture multiple-point change model to make statistical inference about a reference cycle by means of several Monte Carlo experiments that try to capture some basic stylized facts that characterize business cycle fluctuations. Among them, we include double dip recessions, uncertain turning points, short samples of specific turning points and asymmetric cycles, which appear when the disaggregated economic indicators are not roughly coincident. With these simulations, we show that the model exhibits a high performance in determining the number of distinct clusters, and that business cycle inferences and parameter estimates are computed very accurately, even in the presence of these business cycle features.

Finally, we perform an empirical application that illustrates how the method developed in this paper can be used by practitioners seeking to construct NBER-like reference cycles. The disaggregated coincident economic indicators used to determine the reference cycle is the set of ten monthly coincident indicators used by the Business Cycle Committee when determining the trough of the Great Recession in the US. The estimated model, applied to the Bry-Boschan turning points of this set of coincident indicators, is stable in this environment with no evidence of switching or non-convergence problems. The estimated number of distinct cycles, as in the case of the NBER-referenced cycle, is eight, being the means, rather than the variances, the parameters that present the highest ability to classify the draws into the separate cycles. Notably, although the NBER dates represent the informal consensus of the committee members and our method is based on replicable statistical rules, both dates roughly coincide.

In addition, we investigate the pseudo real-time performance of the mixture multiple-point change model for dating the US business cycles over the past sixty years, with particular interest in the ability of the model to accurately and quickly identify in real time the eight distinct NBER peaks and troughs over this period. Our results suggest that the model identifies in real time each of the NBER business cycle episodes, and the resulting pseudo real-time dating is in fairly close proximity to the official dates. The model also provides a substantial improvement over the NBERs business cycle dating committee in the speed with which the turning points are identified, especially when dating the troughs. Overall, the model also tends to show some timing advantages with respect to the dating methods described in Chauvet and Piger (2008) and Hamilton (2011), and, to a lesser extent, in Giusto and Piger (2017).

The rest of the paper is organized as follows. Section 2 describes the methodological considerations when turning point dates are viewed as outcomes of a Markov-switching finite mixture distribution whose phase shifts are restricted as in a multiple change-point model. Section 3 proposes a Monte Carlo experiment to analyze the ability of the proposal to date business cycle turning points. Section 4 examines the ability of this new approach to make inferences about the NBER-referenced US business cycle turning points. Section 5 concludes.

2 Finite Markov mixture distribution

2.1 Baseline model

We assume that the reference cycle of the overall economy over a time span is determined by a reference cycle chronology of K pairs of peak and trough dates. This set of reference turning points allows us to partition the reference cycle into non-overlapping episodes of expansions and recessions, each of which contain a single pair of reference turning point dates, $\mu_k = \left(\mu_k^P, \mu_k^T\right)$, where $\mu_k^P < \mu_k^T < \mu_{k+1}^P$ for all $k = 1, \ldots, K$, that separate the business cycle episodes.[3]

Although they are not observed, we can infer the reference turning points from a set of R economic indicators. Let us collect the specific pairs of turning point dates of each coincident indicator r in sets of size n_r, with $r = 1, ..., R$.[4] This produces a total of $N = n_1 + ... + n_R$ individual pairs of turning points, which are collected in $\tau = (\tau_1, ..., \tau_N)$, where $\tau_i = \left(\tau_i^P, \tau_i^T\right)$, $i = 1, \ldots, N$. In practice, the observed pairs of turning point dates τ_i are stacked in ascending order, meaning that $\tau_i^P \leq \tau_{i+1}^P$ and $\tau_i^T \leq \tau_{i+1}^T$ for $i = 1, \ldots, N-1$.

Burns and Mitchell (1946) stated that the dates of specific turning points are concentrated around the K distinct pairs of reference turning points. Therefore, it is reasonable to assume that the distribution of specific turning points is heterogeneous across and homogeneous within the reference turning points. We deal with these empirical facts by assuming that the turning points arise at each episode k from the same parametric family of bivariate Gaussian distributions, whose different means are the reference turning points, μ_k, and whose different covariance matrices are Σ_k.

When the dates of the episodes are unknown, a natural way to deal with this specification is to assume that the specific turning point dates are realizations of a bivariate random variable arising

[3] If one considers separate episodes of trough-peak dates, the model can be described symmetrically by defining $\mu_k = \left(\mu_k^T, \mu_k^P\right)$. This is discussed in more detail in the pseudo-real time application.

[4] The specific turning point dates could, for example, be the output of a variant of the Bry and Boschan (1971) dating algorithm.

from a mixture of K normal distributions. Define the unknown parameters of the distributions as $\theta = (\theta_1, ..., \theta_K)$, where $\theta_k = (\mu_k, \Sigma_k)$ are the distribution parameters in group k.

The mixture model can be viewed as a hierarchical latent variable model. We assume that the reference cycle turning points may be labeled through an unobservable state variable s taking values in the set $\{1, \ldots, K\}$ in the whole sequence of realizations, which are collected in $S = (s_1, \ldots, s_N)$. Therefore, $s_i = k$ indicates that the observation τ_i is drawn from a bivariate Gaussian distribution with parameters θ_k.

The state variable is modeled as a first-order K-state Markov chain, which implies that the probability of a change in regime depends on the past only through the value of the most recent regime

$$\Pr(s_i = k|s_{i-1} = l, ..., s_1 = w, \tau^{i-1}) = \Pr(s_i = k|s_{i-1} = l) = p_{lk}, \tag{1}$$

where $\tau^i = (\tau_1, ..., \tau_i)$. The fact that both the pairs of reference turning points and the pairs of observed specific turning points are in ascending order implies that $s_i \leq s_{i+1}$.

The stochastic properties of this process are described by a $(K \times\ K)$ transition matrix, P, whose rows sum to one. Within our business cycle framework, it makes sense to consider the regime switches as driven by a nonergodic Markov chain, which captures the switch from one pair of turning point dates to the following pair as a multiple structural break or multiple change-point model. That is, once the Markov chain reaches a reference cycle date k for the first time, the process remains there with probability p_{kk} until it reaches the following reference cycle date $k+1$ with probability $1 - p_{kk}$, for $k = 1, \ldots, K-1$. Once the Markov chain reaches the latest reference cycle date, it remains there with probability one.

Following Chib (1998), the unobserved state variable would exhibit the following transition probability matrix

$$P = \begin{pmatrix} p_{11} & 1-p_{11} & 0 & \cdots & 0 \\ 0 & p_{22} & 1-p_{22} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & & 0 & p_{K-1K-1} & 1-p_{K-1K-1} \\ 0 & \cdots & & 0 & 1 \end{pmatrix}. \tag{2}$$

Note that this parameterization of the transition probability matrix automatically enforces the order constraints on the reference turning point dates. In this case, P depends upon the set of probabilities p_{kk}, with $k = 1, ..., K-1$, that are collected in the vector π.

To sum up, the final problem is assigning each specific turning point to a certain reference cycle turning point by making inference on the unobserved allocations S. For this purpose, each τ_i is assumed to be a realization of a finite Markov mixture model of K components and, from the law of total probability, its marginal density is

$$p\left(\tau_i|\theta, \pi, \tau^{i-1}\right) = \sum_{k=1}^{K} \Pr\left(s_i = k|\theta, \pi, \tau^{i-1}\right) p\left(\tau_i|\theta_k, \tau^{i-1}\right). \tag{3}$$

Here, the mixture weight $\Pr\left(s_i = k|\pi, \tau^{i-1}\right)$ is the filtered probability of sampling from group k, with $0 \leq \Pr\left(s_i = k|\theta, \pi, \tau^{i-1}\right) \leq 1$, and $\sum_{k=1}^{K} \Pr\left(s_i = k|\theta, \pi, \tau^{i-1}\right) = 1$, which is governed by the transition probability matrix P. In our context, $p\left(\tau_i|\theta_k, \tau^{i-1}\right)$ is the Gaussian density, $N\left(\mu_k, \Sigma_k\right)$, with different means and covariance matrices.

2.2 Bayesian estimation using the Gibbs sampler

The estimation of the parameters collected in θ and π and the inference about S is performed through a Markov Chain Monte Carlo (MCMC) method.[5] The Gibbs sampler used to implement the MCMC is started from some preliminary classification $S^{(0)} = \left(s_1^{(0)}, ..., s_N^{(0)}\right)$, which reveals the number of observations assigned to each k-th turning point, $N_k(S^{(0)})$, and its sample means $\mu_k^{(0)}$, with $k = 1, \ldots, K$.[6] The distribution of the parameters can be approximated by the empirical distributions of simulated values by iterating the following steps for $m = 1, ..., M_0, M_0 + 1, ..., M_0 + M$, where the first M_0 are discarded to ensure the convergence of the Gibbs sampler.

STEP 1. Sample the model parameters $\pi^{(m)}$ and $\theta^{(m)}$ conditional on the data and the classification $S^{(m-1)}$. For finite mixture models, it is common to assume that the prior distribution of θ and π are independent, which then also turns out to hold for the posterior distributions conditional on S and the data.

When sampling the transition probabilities π, the Beta distribution is the standard choice in the context of modeling time-series specifications that are subject to multiple change-points. Albert and Chib (1993) show that the transition probabilities may be simulated without regard to the sampling model of the data τ. Following Chib (1998), we assume that p_{kk} are independent a priori, $p\left(p_{11}, ..., p_{K-1,K-1}\right) = p(p_{11}) \cdots p(p_{K-1K-1})$, and follow Beta distributions, $p_{kk} \sim Beta\left(e_{k1}, e_{k2}\right)$. They remain independent a posteriori and also follow Beta distributions $p_{kk}|S \sim Beta\left(e_{k1}\left(S\right), e_{k2}\left(S\right)\right)$, where

$$e_{k1}\left(S\right) = e_{k1} + N_{kk}(S), \quad (4)$$

$$e_{k2}\left(S\right) = e_{k2} + N_{kk+1}(S), \quad (5)$$

for $k = 1, ..., K-1$. In this expression, $N_{ij}(S)$ is the number of transitions from i to j, and, according to the transition probability matrix assumed in (2), $N_{kk+1}(S) = 1$.[7] Then, one can sample $\pi^{(m)}$ from the posterior distributions.

When sampling θ, we propose using a prior which has μ_k coefficients and Σ_k covariances that are independent of one another. When holding the bivariate mean μ_k of each group of turning point dates fixed, under the same conjugate Wishart prior for each (inverse) covariance matrix

[5] For further details on this topic, interested readers are referred to the excellent review by Fruhwirth-Schnatter (2006).

[6] We assume that S^0 is independent of P. In the empirical example, the preliminary classification of turning point dates is the output of a k-means type clustering algorithm.

[7] The probabilities p_{kk} can be simulated by letting $p_{kk} = x_{1k}/\left(x_{1k} + x_{2k}\right)$, where $x_{1k} \sim Gamma\left(e_{k1} + N_{kk}(S), 1\right)$ and $x_{2k} \sim Gamma\left(e_{k2}(S) + 1, 1\right)$.

$\Sigma_k^{-1} \sim W\left(c_{0k}, C_{0k}^{-1}\right)$, the posterior density is $\Sigma_k^{-1}|\mu_k, S, \tau \sim W\left(c_k\left(S\right), C_k\left(S\right)^{-1}\right)$, where

$$c_k\left(S\right) = c_{0k} + N_k(S) \tag{6}$$
$$C_k\left(S\right) = C_{0k} + \sum_{i:s_i=k}\left(\tau_i - \mu_k\right)\left(\tau_i - \mu_k\right)'. \tag{7}$$

Then, sampling $\Sigma_k^{-1(m)}$ given the data, $S^{(m-1)}$, and $\mu_k^{(m-1)}$ is straightforward.

In addition, when maintaining the covariances Σ_k of each group of turning point dates, under the prior for each mean $\mu_k \sim N\left(b_{0k}, B_{0k}\right)$, the posterior density is $\mu_k|\Sigma_k, S, \tau \sim N\left(b_k\left(S\right), B_k\left(S\right)\right)$, where

$$B_k\left(S\right) = \left(B_{0k}^{-1} + N_k(S)\Sigma_k^{-1}\right)^{-1} \tag{8}$$
$$b_k\left(S\right) = B_k\left(S\right)\left(B_{0k}^{-1}b_{0k} + \Sigma_k^{-1}\sum_{i:s_i=k}\tau_i\right). \tag{9}$$

Again, one can easily sample $\mu_k^{(m)}$ given the data, $S^{(m-1)}$ and $\Sigma_k^{-1(m)}$.

A variety of priors can be used in the context of Markov-switching mixture models to implement the Gibbs sampler. In the empirical application, following Chib (1998), we use the same priors for each group of transition probabilities, setting $e_{k1} = 6$ and $e_{k2} = 0.1$.[8] For the independent Normal-Wishart parameters, we work with the noninformative priors $b_{0k} = 0$, $c_{0k} = 0$, $B_{0k} = 1000I_2$, and $C_{0k} = I_2$.[9]

STEP 2. Multi-move sampling of the classification $S^{(m)}$ conditional on the parameters $\pi^{(m)}$ and $\theta^{(m)}$. The multi-move Gibbs procedure is based on sampling the whole path S from its conditional posterior distribution, which, according to the properties of the first-order Markov chain, can be viewed as

$$\Pr\left(S|\theta,\pi,\tau\right) = \Pr\left(s_N|\theta,\pi,\tau^N\right)\prod_{i=1}^{N-1}\Pr\left(s_i|s_{i+1},\theta,\pi,\tau^i\right). \tag{10}$$

Chib (1996) showed that $\Pr\left(s_i|s_{i+1},\theta,\pi,\tau^i\right) \propto \Pr\left(s_i|\theta,\pi,\tau^i\right)\Pr\left(s_{i+1}|s_i\right)$. Thus, there are two ingredients in expression (10), the filtered probabilities and the transition probabilities.

Then, sampling $S^{(m)}$ requires computing the filtered probabilities of each state, $\Pr\left(s_i = k|\theta,\pi,\tau^i\right)$. Hamilton (1989) describes a forward filter to store these filtered probabilities. Starting from an initial value $\Pr\left(s_0 = k|\theta,\pi,\tau^0\right)$, with $k = 1,\ldots,K$, the following two steps are carried out recursively for $i = 1,...,N$.[10] First, the one-step ahead prediction with the information up to turning point $i-1$

$$\Pr\left(s_i = k|\theta,\pi,\tau^{i-1}\right) = \sum_{l=1}^{K}p_{lk}\Pr\left(s_{i-1} = l|\theta,\pi,\tau^{i-1}\right) \tag{11}$$

is computed. Second, when the i-th turning point is added, the filtered probability is updated as

[8] Note that the prior mean duration of each episode k is $\left(e_{k1}+e_{k2}\right)/e_{k2}$. Thus, we are assuming that the prior expected duration of each episode is approximately 60 months in the empirical applications.

[9] Alternative priors appear, for example, in Robert (1996) and Bensmail et al. (1997).

[10] In the applications, we set $Pr\left(s_0 = 1|\theta,\pi,\tau^0\right) = 1$, and $Pr\left(s_0 = k|\theta,\pi,\tau^0\right) = 0$ for $k = 2,...,K$.

follows

$$\Pr\left(s_i = k|\theta, \pi, \tau^i\right) = \frac{\Pr\left(s_i = k|\theta, \pi, \tau^{i-1}\right) p\left(\tau_i|\theta_k, \tau^{i-1}\right)}{p\left(\tau_i|\theta, \pi, \tau^{i-1}\right)}, \tag{12}$$

where $p\left(\tau_i|\theta, \pi, \tau^{i-1}\right) = \sum_{k=1}^{K} \Pr\left(s_i = k|\theta, \pi, \tau^{i-1}\right) p\left(\tau_i|\theta_k, \tau^{i-1}\right)$.

Using the stored filtered probabilities, the states can be simulated from their join distribution (10) starting by sampling the state of the last observation s_N from the smoothed probability $\Pr\left(s_N|\theta, \pi, \tau^N\right)$, which coincides with the last filtered probability. In particular, we generate a random number, u, from a uniform distribution between 0 and 1. In addition, we compute $\sum_{k=1}^{k^*} \Pr\left(s_N = k|\theta, \pi, \tau^N\right)$, $k^* = 1, ..., K$. Then, we call w the number of times that these accumulated probabilities are lower than u. Finally, we sample s_N as $1 + w$.

Now, one can obtain the conditional distribution of the states, $\Pr\left(s_i|s_{i+1}, \theta, \pi, \tau^i\right)$, by backward recursion. For example, if $s_{i+1} = l$, Chib (1996) showed that the conditional distribution of $s_i = j$ is

$$\Pr\left(s_i = j|s_{i+1} = l, \theta, \tau^i\right) = \frac{p_{jl} \Pr\left(s_i = j|\theta, \tau^{i-1}\right)}{\sum_{k=1}^{K} p_{kl} \Pr\left(s_i = k|\theta, \tau^{i-1}\right)}. \tag{13}$$

Again, we generate the states s_i, $i = N-1, N-2, ..., 1$, by comparing the conditional distributions with random numbers generated from a uniform distribution.

As documented by, among others, Frühwirth-Schnatter (2001), the posterior of the mixture model could have $K!$ different modes. Accordingly, the unconstrained MCMC sampler described above could have identifiability problems because it is subject to potential label switching problems. To identify the model, we use the identifiability constraint that the draws must imply a segmentation of the time span into K non-overlapping episodes, i.e., $\mu_k^{P(m)} < \mu_k^{T(m)} < \mu_{k+1}^{P(m)}$ for all $k = 1, \ldots, K$.[11] To ensure that the restrictions apply, we employ the rejection sampling, which is achieved by discarding the draws that do not satisfy the constraints and the sampler is implemented again until the condition is satisfied.

2.3 The number of clusters of turning point dates

We have assumed so far that the number of components, K, was known. However, in practice, one needs to infer the number of groups of specific turning point dates that are cohesive and form a distinct cluster separate from other clusters of specific business cycle turning point dates. Notably, despite the huge effort made in this area, the problem of choosing the number of clusters is still unsolved. So, with the aim of robustness, we follow two different approaches in the empirical analysis.[12]

The first approach, followed by Chib (1998), relies on how Bayesian inference is carried out under model uncertainty. In such settings, a useful way to compare two models with different numbers of clusters K_i and K_j, is to compute twice the natural logarithm of the Bayes factor B_{ij},

[11] Note that the permutation sampler described by Frühwirth-Schnatter (2001) that consists of reordering the values of the draws in order to fulfill the state-identifying restrictions does not apply easily in this multivariate framework.

[12] One informal method for proposing a tentative value for the number of clusters could be based on the visual inspection of the number of modes in the histogram or the kernel density of the data.

which is the ratio of the two integrated likelihoods that correspond to the models with K_i and K_j clusters, respectively. This measure can be approximated by the difference between the two corresponding Bayesian Information Criteria (BIC), which provides a measure of whether the data increases or decreases the odds of a model with K_i clusters relative to a model with K_j clusters. Kass and Raftery (1995) propose that values of $2\log(B_{ij})$ less than 2 correspond to weak evidence in favor of the model with K_j clusters, values between 2 and 6 to positive evidence, between 6 and 10 to strong evidence, and greater than 10 to very strong evidence.

Within this framework, we can easily select the number of clusters by applying Bayes factor comparisons sequentially. First, we check whether a model with 2 clusters increases the odds of a model with 1 cluster. If this happens, then we check whether a model with 3 clusters increases the odds of a model with 2 clusters. The sequential procedure stops when a model with $K+1$ clusters does not increase the odds of a model with K clusters. When this happens, the proposed number of clusters is K.

The second approach follows the lines proposed by Koop and potter (2007), who develop a multiple change-point model with an unknown number of breaks, which is taken as random and estimated using the data. In contrast to assuming constant transition probabilities, which implicitly implies that the regime duration would follow a geometric distributions as in Chib (1998), they model the duration of a regime k, d_k, using a Poisson distribution with mean λ_k, $d_k - 1 \sim Po(\lambda_k)$.

The non-constant transition probability from regime k to regime $k+1$ depends on the current duration of regime. Assuming that regime 1 starts with the initial observation, the transition probabilities are

$$\Pr\left[s_{t+1} = k+1 | s_t = k, d_k\right] = \frac{\exp(-\lambda_k)\lambda_k^{d_k-1}}{(d_k-1)!\left(1 - \sum_{j=0}^{d_k-2} \frac{\exp(-\lambda_k)\lambda_k^j}{-j!}\right)}, \tag{14}$$

where the sum in the denominator is defined to be 0 when $d_k = 1$.

These authors propose a hierarchical prior for λ_k such that $p(\lambda_1, \ldots, \lambda_T) = p(\lambda_1)\cdots p(\lambda_T)$ with the prior

$$\lambda_k | \beta_\lambda \sim G\left(\alpha_\lambda, \beta_\lambda\right), \tag{15}$$

where β_λ, which reflects the degree of dissimilarity of the durations, is an unknown parameter, with prior

$$\beta_\lambda^{-1} \sim G\left(\xi_1, 1/\xi_2\right). \tag{16}$$

Then, the conditional posterior for λ_k is

$$\lambda_k | \beta_\lambda \sim G\left(\alpha_k, \beta_k\right), \tag{17}$$

where $\alpha_k = \alpha_\lambda + d_k$ and $\beta_k = \left[\beta_\lambda^{-1} + 1\right]^{-1}$.[13] As discussed in Koop and Potter (2007), Chib's algorithm can still be applied in the case of a non-constant homogeneous transition matrix.

In our context, we adapt the Poisson hierarchical approach for durations suggested by Koop

[13] Koop and Potter (2007) describe a simple accept/reject algorithm to draw λ_k for the last regime where the duration is not observed.

and Potter (2007) to determine a tentative number of clusters rather than to estimate the model's parameters for two reasons. First, because the average of the posterior means for the estimates of reference cycle turning points, $\{\mu_1, \ldots, \mu_K\}$, would imply smoothed dates. However, by definition, they are single points in time referring to turning point dates, which are constant within each cycle and change sharply to the next regime at each break date. Second, because the resulting estimated number of regimes must be a positive integer rather than a real number.

Our recommendation is rounding the estimated number of breaks to the nearest positive integer. Then, the Bayes factor comparison described above can be used to determine the number of phases of the reference cycle K around this integer. Finally, the estimates of the finite Markov mixture models for K will determine the single dates of the reference cycle turning points, which break the time span of interest into separated segments that determine the distinct business cycle phases.

2.4 Data problems

The application of finite Markov mixture models to turning point data presents two data challenges. The first is that finite mixtures of Gaussian distributions are defined for continuous data. However, the turning point dates are obtained from coincident indicators that are usually sampled at monthly or quarterly frequencies and this generates a discontinuity challenge.

For example, when the coincident indicators are sampled at monthly frequencies, turning points are dates of format $YYYY.mm$, which refers to month mm of year $YYYY$. In this case, the distance between the eleventh month and the twelfth month of a year is lower than that betwen the twelfth month of the same year and the first month of the following year. To overcome this drawback, we propose the transformation of the turning points to dates $YYYY.d$, where $d = 1/12(mm-1)$, which can obviously be changed back to recover the results in the standard monthly calendar dates.[14]

The second challenge of this approach is that the individual turning points are collected from sets of disaggregated economic indicators and some of them might not be available for the full span. In practice, some indicators are reported with lags exhibiting missing observations in the latest months while others are available only for a diminished time span because they are not available at the beginning of the sample. In our context, this is rarely a problem since it only implies that the probability distributions of turning points at early and end dates must be estimated from a relatively lower number of observations due to missing data.

3 Simulation analysis

In this section, we set up several experiments to analyze the performance of the mixture multiple change-point model to determine the correct number of distinct clusters of specific turning point dates, to classify each specific turning point in its correct cluster, and to estimate the reference cycle dates and their covariances. The experiments are designed to guide practitioners as to whether our proposal works as a dating method of reference cycles of regions with business cycle characteristics similar to those that have largely been examined in the related literature.

[14] In a similar fashion, when the coincident indicators are sampled at quarterly frequency, the dates $YYYY.q$ are transformed into dates $YYYY.d$, where $d = 1/4(q-1)$.

3.1 Experiment design

We start the experiment by simulating a basic scenario, called Case 0, in which we simulate specific turning point dates from mixtures of three bivariate Gaussian distributions. For each of these mixtures, the specific dates are generated around three pairs of reference cycle dates, the means of the Gaussian distributions, that agree with the average characteristics of the NBER-designated reference cycle in a monthly calendar. Without loss of generality, we set the first reference peak, μ_1^P, to 1980.01. In accordance with the average duration of the NBER expansions and recessions, the distance between each reference peak and the following reference trough is set to 11.1 months and the distance between each reference trough and the following reference peak is set to 58.4 months.

For each of these three pairs of reference dates, we sample 300 specific dates from bivariate Gaussians distributions whose means, μ_k, are the reference dates and whose covariance matrices, Σ_k, are the average of the covariance matrices that are estimated when the model is applied to the ten monthly US coincident indicators that are the basis of the section devoted to the empirical application.[15] Then, we generate 2500 draws from the posterior distributions although we discard the first 500 to mitigate the effect of initial conditions. This simulation process is repeated 1000 times and we collect the average estimates.

Apart from Case 0, we evaluate the efficacy of the mixture multiple change-point model to make inference about the reference cycle by including simulations that capture some other empirical regularities in business cycle analysis. According to the NBER analysis of the US business cycle, a typical reference cycle may contain unusually long recessions like the Great Recession, which was the longest since the Great Depression. To account for this possibility, in the same fashion as in Case 0, Case 1 simulates a reference cycle whose last recession lasts 18 months. In addition, some recessions are followed by a brief expansion, known as double dip recessions, such as the two dated by the NBER in the eighties. This case is labeled in the simulations as Case 2, whose reference cycle includes an expansion lasting only 12 months.

The effect of determining the specific dates from a set of low quality coincident indicators is examined in Cases 3, 4 and 5. Case 3 tries to capture the scenario where the reference turning point dates are hard to identify because their clusters of specific turning point dates are disperse, that is, they are widely scattered. The simulations in Case 3 use a data generating process whose variances are twice as large as the average of the estimated variances of the empirical application. Case 4 simulates the situation where some indicators are not available at the beginning of the sample, which implies computing inference about the first clusters from a smaller set of observations. In this case, we generate only 50 observations for the first cluster. Case 5 accounts for the scenario where the specific dates are obtained from a small set of indicators, so the total number of simulated specific dates is 50 in this case.

Cases 0 to 5 assume that we are able to collect coincident indicators of the aggregate business cycle. However, in practice, some indicators might have lead or lag patterns for some clusters, which will result in skewness around these clusters of turning points. To examine the effect of departures from the symmetry implied by the Gaussian mixtures, in Cases 6a and 6b, we simulate

[15] To facilitate computations, the covariances are set to zero.

the specific turning points from right-skewed and left-skewed distributions, respectively, by computing a mixture of two distributions. The first one is a normal distribution with the true mean parameters and the second one is a normal distribution that displaces the mean parameters by -3 months in Case 6a and by +3 months in Case 6b. The proportion of the two distributions is 90% and 10%, respectively.

3.2 Model selection

We examine the accuracy of the mixture multiple change-point model approach in grouping the simulated data set of pairs of specific turning point dates in the correct number of clusters in Table 1. For each case, Panel A reports the average over the 1000 simulations of the mean and the median values for the estimated number of distinct clusters from the Poisson hierarchical approach for duration developed by Koop and potter (2007). For all cases, the average number of estimated breaks are between 2.67 and 2.89, suggesting three separate clusters of turning points.

Panel B shows twice the log of the Bayes factor for models with K versus $K-1$ clusters. For each Bayes factor comparison, the figures in brackets show the ratio of the number of times, given as a fraction of unity, that the model with K clusters is achieved over the 1000 simulations, which estimates the probability that the model selection procedure chooses K clusters in the mixture model. Regardless of the simulated case, the sequence of averaged (twice the log of) Bayes factors points to $K = 3$ because the differences of the averaged BICs are well above the values that indicate strong evidence for the model with larger number of clusters for the comparisons of the models with $K = 1$ versus $K = 2$ and the models with $K = 2$ versus $K = 3$. However, increasing K from 3 to 4 clusters implies that the difference of the averaged BICs of the models with $K = 3$ versus $K = 4$ is always negative, which decreases the averaged odds of a model with $K = 4$ clusters relative to a model with $K = 3$ clusters.

For completeness, Panel C shows the averages of some likelihood-based measures for models with $K = 1, ..., K^* = 4$ clusters, such as the marginal likelihood (L_K), the Akaike (AIC_K) criterion, the Bayesian Information Criterion (BIC_K), the entropy (EN_K), and the BIC corrected for misclassification ($BIC_K + EN_K$).[16] In addition, in brackets, the table shows the ratio of the number of times that the model selection procedure chooses model with K clusters over the total number of simulations.

The table shows that these measures unambiguously indicate that the synthetic reference cycle generated in this simulation contains three separate clusters of pairs of specific turning point dates for each of the cases that we consider in the simulation, even if long recessions, short expansions, low quality of indicators, low number of indicators, or non-coincident indicators were involved in the dating procedure. Notably, regardless of the model selection procedure, the ratios reported in brackets show that the mixture model is estimated with the right number of three clusters in all the simulations.[17]

[16] The entropy takes the value of 0 for a perfect partition of the data and a positive number that increases as the quality of the classification deteriorates.

[17] The only exception is the marginal likelihood, which overestimate the number of clusters 18% of the time because this method tends to choose models with a large number of components.

3.3 Model classification

In addition to model selection, it is worth examining the ability of the Markov-switching mixture model to provide a posterior classification of the simulated pairs of specific turning point dates into the different time partitions provided by the reference dates. For this purpose, for each of the 1000 simulations, we estimate the posterior classification probabilities from the 2000 MCMC draws by the corresponding relative frequency from the retained state draws

$$\Pr\left(s_i = k|\theta\right) = \frac{1}{M} \neq \left\{s_i^{(m)} = k\right\}, \tag{18}$$

where $\neq \left\{s_i^{(m)} = k\right\}$ counts the number of times that observation i is allocated to cluster k across the M replications, with $k = 1, \ldots, K$. A natural classification indicator $S^K = \left(s_1^K, \ldots, s_N^K\right)$ that arises from the classification probabilities of a Markov-switching mixture model with K clusters will associate each observation τ_i with component k, i.e. $s_i^K = k$, if $Pr\left(s_i = k|\theta\right)$ reaches a maximum at k over the K components.

Let $S^* = \left(s_1^*, \ldots, s_N^*\right)$ be the true allocation of the simulated data. To measure the ability of a mixture model to assign any specific pair of turning points to the kth group with the classification indicator S^K when it actually arises from that group according to the true classification S^*, we use the classification rate

$$CR\left(S^K, S^*\right) = \frac{1}{N}\sum_{i=1}^{N} I_{\{s_i^K = s_i^*\}}, \tag{19}$$

where $I_{\{s_i^K = s_i^*\}}$ takes the value of one whenever observation i is classified by S^K in the same cluster as the true allocation S^* and zero when this observation is misclassified. Evidently, an allocation S^K with a classification rate of 1 indicates perfect classification because $s_i^K = s_i^*$ for all $i = 1, \ldots, N$.

For each case, the extent to which the selection of the number of clusters in the mixture model of simulated pairs of turning point dates determines a good partition of the underlying reference date is examined in the last panel of Table 1. The table reports the average value over the 1000 simulations of the classification rate for various numbers of clusters. For example, in Case 0, when the number of clusters is under-fitted, which occurs when $K = 2$, only 46% of the observations are allocated in the correct cluster. The classification rate of the model with an over-fitted number of clusters $K = 4$ is 0.5, indicating that about half of the observations are misclassified. Notably, the classification rate of the model that correctly uses $K = 3$ clusters is one, which implies that, on average, all the observations are allocated correctly.

Although we used Case 0 as a basis to describe the models' classification accuracy, a detailed inspection of this table reveals that all the results obtained for this case are consistent for each of the cases that we consider in the simulation study. This highlights the valuable ability of the model to provide an accurate classification of the specific pairs of turning point dates into the different phases of the reference cycle even when the latter contains double dip recessions, uncertain turning points, short samples of specific turning points, and asymmetric cycles, which appear when the disaggregated economic indicators are not coincident.

3.4 Model estimation

In this section, we use the MCMC draws for each of the 1000 simulations for the analysis of the accuracy of the mixture multiple change-point model to recover the true parameters of the means and the variances of the K distributions of the mixture.

To check for the accuracy of the model to estimate the means, the first panel of Table 2 shows the Mean Squared Errors (MSE) as the averages over the 1000 simulations of the squared differences between the posterior means and the true reference dates. Compared to the values of the true reference dates, which range from 1980 to 2000, the estimated MSE, which range from 1.50 to 10.17, indicate that the estimated parameter values tend to be very close to the true parameter values. This high precision in parameter estimation is also observed in the low MSE in the comparison of the posterior variances and the traces of the covariance matrices and their respective true values.

Notably, these accurate results are consistent regardless of the case being evaluated, even in the cases where inference about the reference cycle is computed from a reduced set of coincident indicators or from a set of low-quality indicators. However, the accuracy in recovering the true parameters diminishes a bit when the specific reference dates are computed from non-coincident indicators.

4 Empirical results

The objective of this section is to provide an empirical assessment of the extent to which the mixture multiple change-point model described in this paper is able to compute accurate inferences about the reference business cycle turning points in an empirical example with real data. For this purpose, we focus on the comparison of chronologies of two methods of dating the US reference cycle: the NBER-referenced business cycle dates, which serves as the standard for accuracy, and the reference dates obtained from the mixture multiple change-point model.

4.1 Dating the historical turning points

The set of disaggregated coincident economic indicators that we use to evaluate the method is based on the latest available decisions of the NBER Business Cycle Dating Committee about the timing of the US business cycle turning points. In its most recent memorandum explaining the June 2009 trough (NBER Business Cycle Dating Committee 2010), the Committee mentions that it pays particular attention to ten monthly indicators when determining the trough of the Great Recession in the US reference cycle.[18] These indicators are a measure of monthly GDP that has been developed by the private forecasting firm Macroeconomic Advisers, three measures of monthly GDP and GDI that have been developed by Stock and Watson (2010b), real manufacturing and trade sales, industrial production, real personal income excluding transfers, the payroll and household

[18] An Excel spreadsheet containing the data for the indicators of economic activity considered by the committee is available at

mirror.nber.org/cycles/BCDCFiguresData100920_ver5.xls measures of total employment, and an aggregate of hours of work in the total economy. Although the number of available observations differs across indicators due to different starting dates and different release lags, the largest sample spans the period from January 1959 to August 2010, during which there were eight complete NBER-referenced business cycles.[19]

The first step of the empirical illustration consists of identifying the individual chronologies of turning points in each of the ten indicators. To this end, we apply the peak and trough dating algorithm code implemented by Watson (1994), who followed the lines suggested by Bry and Boschan (1971).[20] Figure 1 provides a preliminary inspection of the data by plotting the kernel density estimator of the bivariate distribution of the resulting pairs of specific peaks and troughs. The figure reveals that the distribution of turning points is multimodal, exhibiting various modes that cluster the individual turning points around periods of NBER-referenced recoveries and declines. The modes of the kernel density of pairs of specific peaks and troughs suggest that the tentative number of different clusters in the reference cycle could be eight, which correspond with the distinct local maxima.

To determine the number of clusters formally, we compute sequentially (twice the log of) the Bayes factors that compare two models with $K-1$ and K different numbers of clusters until we obtain the first value of K for which an additional cluster does not increases the odds of the model with less clusters. This method points to $K=8$ because the differences of the BICs are well above the values that indicate evidence for the models with larger number of clusters for the comparisons of the models from $K=1$ versus $K=2$ to $K=7$ versus $K=8$. However, increasing K from 8 to 9 clusters implies that (twice the log of) the Bayes factor of the models with $K=8$ versus $K=9$ turns to negative (-22.35), which decreases the odds of a model with $K=9$ clusters relative to a model with $K=8$ clusters. This supports the conclusion that there is strong empirical evidence that the number of complete cycles in the US reference cycle from 1959 to 2010 is eight, as the NBER-referenced chronology establishes.

To reinforce this result, we adapt the Poisson hierarchical approach for durations suggested by Koop and Potter (2007) to estimate the number of clusters. Using this approach to our turning point data, we find that the posterior mean of the number of clusters is 7.96 (with posterior median of 8 clusters). This also suggests that the number of distinct cycles is eight, reinforcing the result of the sequential Bayes factor comparison.

For completeness, Table 3 also reports some measures that are typically used in the context of mixture models to perform model selection. The first column of the table shows that the log of the marginal likelihood increases uninterruptedly when the number of clusters increases from $K=1$ to $K=9$, where it reaches a peak at -53.48. This corresponds with $K=9$, although this value is very close to that obtained for $K=8$, -54.33. Although this suggests choosing $K=9$, the marginal likelihood does not take into account the number of parameters in model selection and tends to overestimate the number of clusters.

Regarding model selection criteria that introduce penalties in model selection, the reported

[19] The Great Recession of 2008 marked the last complete cycle in our sample. Therefore, no further specific turning point is detected since then and the sample is still valid to compute historical dates.

[20] In short, the algorithm isolates local minima and maxima of each of the ten indicators, subject to constraints on both the length and amplitude of expansions and contractions.

values in Table 3 of AIC, BIC and BIC corrected by misclassification reach their minimums of 202.66, 297.01, and 297.01, respectively, when $K = 8$. This indicates that the US reference cycle requires eight separate cycles. In addition, the entropy of the mixture model with eight clusters is zero, showing that the model produces a clear segmentation of the reference cycle.

The methods used to choose the number of clusters do not reveal which are the parameters that determine the formation of the different clusters and do not evaluate the extent to which the mixture model provides a good partition of the time span as the NBER chronology establishes. To examine these questions, we consider sampling representations from the rejection sampler described in Section 2, which is a very useful tool for visualizing the posterior mixture distribution. For each of the eight clusters, Figure 2 displays the two-dimensional scatter plots of the MCMC draws for $\left(\mu_k^P, \mu_k^T\right)$ in Panel A and for $\left(\mu_k^P, \Sigma_{11,k}\right)$ and $\left(\mu_k^T, \Sigma_{22,k}\right)$ in Panel B. From these scatter plots, it is obvious that the component parameters differ mainly in the means, which present the highest ability to divide the draws into the eight separate groups, whereas the variances are quite similar for many groups. In addition, Panel A suggests that the clusters of turning points alternate sequentially, supporting the nonergodic behavior imposed in the transition matrix to capture the transitions across the distinct business cycles, as a multiple-change model, which roughly occur about at the NBER-designated turning points.

Figure 3 shows diagnostic plots of the post burn-in draws from the conditional distributions of μ_k^P, μ_k^T, $\log\left(|\Sigma_k|\right)$ and $p_{k,k}$ and $p_{k,k+1}$ for each of the $K = 8$ clusters. These panels help us to detect potential convergence problems or label switching, which arise when the mixture likelihood function is not invariant to relabeling the components of a mixture model. The paths of the draws show that the rejection sampler that imposes, $\mu_k^{P(m)} < \mu_k^{T(m)} < \mu_{k+1}^{P(m)}$, for all $k = 1, \ldots, K$, is useful to prevent label switching because the sampler stays within the modal regions corresponding to the initial labeling. So, these regions are well separated from the others leading to a unique labeling.

Table 4 evaluates the mixture multiple change-point model in terms of its capacity to determine the NBER reference cycle dates. The columns labeled as NBER and MSMM report the reference cycle dates as determined by the NBER and our Markov-Switching Mixture Model. The distinct means are estimated from the posterior distributions with the help of the rejection Gibbs sampler algorithm for the mixture model. Using the outputs of the MCMC algorithm, this table also reports, in brackets, the range of values of the posterior probability distributions that includes 95% of the probability.[21] It is evident that the mixture model replicates the NBER peak and trough dates very accurately. The method provides exact matches for one peak and three troughs, differing by less than one quarter either way in the date of the remaining turning points.[22] On average, the method locates the peaks about one and a half months before and the troughs about one sixth of a month after the NBER dating. Remarkably, there are no instances in which an NBER turning point date is not matched by a similar date produced by the MSMM model, whose 95% credible intervals include the NBER-referenced dates in all the episodes.

[21] The credible intervals, which range from 0.08 to 1.08 years for peaks and from 0.07 to 1.98 years for troughs, provide a precise estimate of the reference dates.

[22] The only exception is the 2001 trough, which is located only four months after the NBER dating. The difficulties of identifying this mild recession have been documented in Kliesen (2003) and Hamilton (2011).

To provide the results with additional economic meaning, it is worth examining the ability of the mixture multiple change-point model to provide a posterior classification of the specific turning point dates into the different business cycles of the US reference cycle. With this aim, we estimate the posterior classification probabilities $\Pr(s_i = k|\theta)$, with $k = 1, \ldots, 8$ and $i = 1, \ldots, N$, from the MCMC draws through the corresponding relative frequency of the retained state draws, as described in Section 3. Figure 4, which plots the estimated classification probabilities for each cluster, shows that the model clearly classifies each specific date, whose classification probabilities are always close to either 1 or 0. In addition, the probabilities agree with the multiple change-point behavior of the model, with a large persistence of each state and probabilities that increase nearly to unity after each structural break that occurs about the turning point dates identified by the NBER. Thus, the probabilities support the view that the mixture multiple change-point model performs a segmentation of the time span into non-overlapping business cycle episodes that agrees with the NBER-referenced cycles.

4.2 Pseudo real-time performance

Although the ex-post identification of the reference cycle turning points is of great interest in itself, doing this on an out-of-sample basis is a bigger challenge (Hamilton, 2011). In this section, we evaluate both the accuracy and the speed with which the mixture multiple change-point model would have dated the US reference cycle as additional specific cycle turning points were appearing in (pseudo) real time. Again, the NBER chronology is the basis of comparison.

For this purpose, we use the same data of ten coincident economic indicators described in the in-sample analysis, which are available from 1959.01 to 2010.08. However, we develop here a pseudo real-time analysis that consists of computing inferences from successive partitions (one-month enlargements) of the latest available data set.[23] In short, the new information provided by the ten economic indicators is incorporated into the model month by month as we enlarge the data set and the model converts the turning point detection into a classification problem. Each month, the model determines whether the set of incoming specific turning points dates generates a separate cluster, which would imply that a phase change occurs. To avoid false signals, we follow Chauvet and Piger (2008) and require that the Bayes factor indicates an increase in the number of clusters for 3 consecutive months, to confirm that a new reference cycle turning point has appeared. In addition, since the reference cycle dates are the means of the clusters of the specific dates, we do not date the new reference turning point until the new specific turning points appear in at least one-third of the macroeconomic indicators.[24]

Table 5 evaluates the capacity of the mixture multiple change-point model to date the US reference cycle in pseudo real time. For a comparative assessment, the left panel reports the NBER peaks and troughs and the months of their respective announcements, while the right panel displays the estimated reference cycle turning point chronology produced by the mixture multiple change-point model and the months when the model detects the phase changes. Using the sample of the

[23] Producing real-time vintages is unfeasible because the historical records of some of the ten indicators are not available.

[24] Estimating the means with only the early available specific turning points would erroneously produce ex-ante estimates lower than the ex-post estimates.

posterior draws, the table reports not only the sample means as point estimates but also the credible intervals, which are constructed using the 2.5% and 97.5% sample quantiles. Then, comparing the NBER dates with those established by MSMM measures the accuracy, while comparing the official months of the announcements with the months in which MSMM identifies the phase changes quantifies the speed of detection.

The specification whose clusters are determined by Gaussian distributions with means $\left(\mu_k^P, \mu_k^T\right)$, the peaks and troughs of the reference cycle, is called peak-trough MSMM. On the other hand, the model that determines a reference cycle beginning with a trough and whose specific troughs and peaks are clustered around the means of bivariate Gaussians distributions $\left(\mu_k^T, \mu_k^P\right)$, is called trough-peak MSMM. The recursive analysis starts with a vintage of data that covers the period 1959.01-1966.01, which includes one pair of NBER peak-trough dates. Using this sample, we identify the specific turning points for each indicator by implementing the Bry-Boschan dating algorithm. Then, we employ the Bayes factor of the peak-trough MSMM model to confirm that there is only one cluster of specific turning point dates, which is generated from a bivariate normal density whose means are the first pair of peak-trough reference dates. According to the figures reported in Figure 5, μ_1^P and μ_1^T are dated in 1960.03 and 1961.02, respectively.

Now, the sample is enlarged by adding one month of observations to each of the ten indicators and the Bry-Boschan algorithm is re-estimated for the period 1959.01-1966.02. If the outcome of the dating algorithm does not include any new specific peak, the vintage of data is enlarged again with the observations of a new month. On the contrary, when a new specific peak is detected by the Bry-Boschan algorithm, we follow the simple automatic procedure of creating its artificial pair (trough) by adding the average duration of the preceding recessions to that peak (the preliminary trough will be replaced by the actual trough as soon as it is detected by Bry-Boschan when new data arrive). In this way, we increase the sample month by month and estimate the peak-trough MSMM model with the enlarged outcomes of the Bry-Boschan algorithm. The recursive updated estimation of the first cluster stops when the Bayes factor that compares the model with $K = 1$ versus the model with $K = 2$ indicates that a new cluster has appeared. As Table 5 reports, this happens in 1971.06 and the peak μ_2^P is dated in 1969.11.

The procedure described above does not guarantee an accurate estimation of the trough μ_2^T because it could be partially based on artificial specific troughs. The approach we follow to detect the following trough is to move to the trough-peak MSMM model as soon as a new peak is detected. In this case, we omit the peak dated in 1960.03, which converts the reference cycle in a sequence of clusters centered in pairs of troughs and peaks. Then, the procedure follows the same rules described above: (i) the sample is updated by one period; (ii) the Bry-Boschan routine is used to detect new troughs, which are matched up with artificial peaks that guarantee that the current expansion will last the average of the preceding expansions (replaced by actual peaks when they appear as the outcome of the Bry-Boschan algorithm); and (iii) the peak-trough MSMM model is estimated and the Bayes factor that compares the trough-peak MSMM model with $K = 1$ versus the model with $K = 2$ is computed. The automatic procedure continues until $K = 2$ is preferred to $K = 1$, which happens in 1971.06 as reported in Table 5. At this point, we date the trough μ_2^T in 1970.11 and we move again to the trough-peak MSMM model with the aim of detecting a new phase change and dating the following peak μ_3^P.

This automatic procedure goes back and forth from the peak-trough model to the trough-peak model when new turning points are detected, increasing the sample month by month and applying the classification procedure provided by MSMM to the enlarged data sets until the end of the sample in 2010.08. Table 5 examines the success of an analyst who applies MMSM to date the NBER reference cycle turning point dates in (pseudo) real time each month between 1966.01 and 2010.08, both in terms of timeliness and accuracy. Beginning with the accuracy, the dates assigned to the reference cycle turning points in pseudo real time are in close agreement to those determined by the NBER, although the precision of these estimates varies across episodes. The mixture model produces the same turning points as the NBER in one out of eight peaks and in five out of eight troughs, and are within a quarter of each other in all but one case. The noticeable exception is the peak dated in 1989.12 by our method, well before the peak dated in 1990.07 by the NBER, suggesting that this peak is associated with a leading pattern of some of the coincident indicators. It is remarkable that the credible intervals contain the NBER dates in all cases.

The following question of interest is the timeliness with which this accurate identification is achieved. For this purpose, in Table 5 we report the number of months after the NBER turning point dates that the NBER and MSMM require to identify these turning points. The figures show a systematic and significant improvement over the NBER in the speed with which business cycle turning points are identified. On average, MSMM allowed the peaks to be identified 4.4 months earlier than NBER, and this improvement increases to 11.2 months in the case of troughs. This improvement over the NBER's business cycle dating committee timeliness when dating the troughs has been documented by Chauvet and Piger (2008). Although we were precluded from using true real time data as they did, we find that MSMM tends to improve the timeliness of their dating proposals. In addition, MSMM also provides a substantial improvement in the speed with which the dating methods surveyed by Hamilton (2011) identify the turning points of the Great Recession. Finally, our approach identifies peaks more quickly for 2 of the 5 out-of-sample peaks and for 2 of the 5 out-of-sample troughs than the turning point dates provided by the machine-learning algorithm proposed by Giusto and Piger (2017).

5 Conclusion

Since the early work of Burns and Mitchell (1946), the reference cycle is viewed as a sequence of business cycle fluctuations occurring at about the same turning points in many economic activities, where the dates of the shifts determine the reference dates. However, the fact that the cyclical turns of different processes are concentrated around certain points in time does not imply that dating the reference cycle is merely a matter of counting the economic time series that rise and that fall at about the same time. In fact, dating the reference cycle has been the source of much debate in both the research literature and policy. The approach pursued in this paper to date the reference cycle is based on aggregating the specific turning points from a set of coincident economic indicators.

Our novel proposal can be implemented making only a few distributional assumptions. The set of specific turning points dated from a set of coincident indicators is viewed as a sequence of data with a natural ordering which can be broken down into segments by the reference cycle turning points. In particular, the specific turning points are supposed to be drawn from the same Gaussian distribution within each segment, but from different Gaussian distributions in different segments. Thus, the set of bivariate individual turning points are considered to arise from a mixture distribution where the different model parameters are assumed to evolve according to a latent random discrete-state Markov process with the transition probabilities constrained as in the multiple-change break point model proposed by Chib (1998). In this representation, the means of the different Gaussian distributions are viewed as natural estimates of the reference cycle turning points and the state probabilities are used to determine the change points of the different reference cycles and to classify the specific turning point dates into the different business cycles of the reference cycle. In this context, standard Bayesian estimation of finite Markov mixture modeling techniques is suggested to estimate the model.

The reliability of the proposed framework is validated with simulated data in a Monte Carlo experiment that try to mimic some stylized facts of business cycle analyses such as short business cycle phases, uncertain turning points, short samples and specific cycles that come from non-coincident indicators. The results suggest that the method is very accurate to determine the number of clusters, to classify the set of specific turning points and to recover the true parameters, regardless of the stylized business cycle fact being analyzed.

Finally, although the mixture multiple change-point model determines the reference cycle dates from a set of individual coincident indicators without prior knowledge of any pre-specified dating, it has been evaluated in terms of its capacity to generate the NBER reference cycle dates. For this purpose, we use the list of ten coincident indicators employed by the NBER Business Cycle Committee to determine the trough of the Great Recession. A first benefit of using our method to determine the historical reference cycle is its replicability. In contrast to the committee approach, which largely relies on judgment, our automatic method is simple and easy to replicate. We find that the differences between the ex-post NBER-referenced turning points and the reference cycle estimated with the mixture multiple change-point model that we propose are almost negligible. So, it can be used with the aim of performing an ex-post dating the reference cycle of another region or time.

A second benefit of using our proposal to determine business cycle phase changes is timeliness because the dating committee announces the turning points with a considerable lag. For example, the committee announced the December 2007 peak in December 2008 and the June 2009 trough in September 2010. We evaluate the (pseudo) real-time ability of our proposed algorithm to identify business cycle turning points in the US economy accurately and in a timely fashion. We find that our method identifies the dates of the NBER-references turning points very accurately and with a remarkable speed. Thus, the proposed framework is also a very promising tool for ex-ante detection and estimation of business cycle turning points.

References

[1] Albert, J., and Chib, S. 1993. Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts. *Journal of Business and Economic Statistics* 11: 1-15.

[2] Bensmail, H., Celeux, A., Raftery, E., and Robert, C. 1997. Inference in model-based cluster analysis. *Statistics and Computing* 7: 1-10.

[3] Burns, A., Mitchell, W. 1946. Measuring Business Cycles. New York: National Bureau of Economic Research.

[4] Bry, G., and Boschan, Ch. 1971. Cyclical Analysis of Time Series: Procedures and Computer Programs. New York: National Bureau of Economic Research.

[5] Business Cycle Dating Committee. 2010. Determination of the June 2009 trough in economic activity. National Bureau of Economic Research. http://www.nber.org/cycles/sept2010.html.

[6] Chauvet, M., and Piger, J. 2008. A comparison of the real-time performance of business cycle dating methods. *Journal of Business and Economic Statistics* 26: 42-49.

[7] Chib, S. 1996. Calculating posterior distribution and model estimates in Markov mixture models. *Journal of Econometrics* 68: 339-360.

[8] Chib, S. 1998. Estimation and comparison of multiple change-point models. *Journal of Econometrics* 86: 221-241.

[9] Fraly, Ch. and Raftery, E. 2002. Model-Based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association* 97:458, 611-631.

[10] Frühwirth-Schnatter, S. 2001. Markov Chain Monte Carlo estimation of classical and dynamic switching and mixture models. *Journal of the American Statistical Association* 96: 194-209.

[11] Frühwirth-Schnatter, S. 2004. Estimating marginal likelihoods for mixture and Markov switching models using bridge sampling techniques. *The Econometrics Journal* 7: 143-167.

[12] Frühwirth-Schnatter, S. 2006. Finite Mixture and Markov Switching models. Springer Series in Statistics. New York, NY: Springer.

[13] Giusto, A., and Piger, J. 2017. Identifying business cycle turning points in real time with vector quantization. *International Journal of Forecasting* 33: 174-184.

[14] Hamilton, J. 1989. A new approach to the economic analysis of nonstationary time series and the business cycles. *Econometrica* 57: 357-384.

[15] Hamilton, J. 2011. Calling recessions in real time. *International Journal of Forecasting* 27: 1006-1026.

[16] Harding, D., and Pagan, A. 2006. Synchronization of cycles. *Journal of Econometrics* 132: 59-79.

[17] Harding, D., and Pagan, A. 2016. The econometric analysis of recurrent events in macroeconomics and finance. New Jersey: Princeton University Press.

[18] Kass, R., and Raftery, A. 1995. Bayes factors. *Journal of the American Statistical Association* 90: 773-795.

[19] Kliesen, K. 2003. The 2001 recession: How was it different and what developments may have caused it? *Federal Reserve Bank of St. Louis Review* September/October: 23-37.

[20] Koop, G., and Potter, S. 2007. Estimation and forecasting in models with multiple breaks. *Review of Economic Studies* 74: 763-789.

[21] Robert, C. 1996. Mixtures of distributions: Inference and estimation. In W. Gilks, S. Richardson, and D. Spiegelhalter (eds.), *Markov Chain Monte Carlo in practice.* London: Chapman and Hall.

[22] Stock, J., and Watson, M. 2010a. Indicators for dating business cycles: cross-history selection and comparisons. *American Economic Review: Papers and Proceedings* 100: 16-19.

[23] Stock, J., and Watson, M. 2010b. New Indexes of Monthly GDP, available at *http://www.princeton.edu/˜ mwatson/mgdp_gdi.html.*

[24] Stock, J., and Watson, M. 2014. Estimating turning points using large data sets. *Journal of Econometrics* 178: 368-381.

[25] Watson, M. 1994. Business-cycle durations and postwar stabilization of the US economy. *American Economic Review* 84: 24-46.

Table 1: Model selection and model classification

	Case 0				Case 1			
	Panel A. Poisson hierarchical approach for durations							
	Mean: 2.88		Median: 2.89		Mean: 2.84		Median: 2.84	
K	1	2	3	4	1	2	3	4
	Panel B. Based on Bayes factor							
$K-1/K$	-	1979.98	2264.92	-43.88	-	2158.80	2326.06	-46.38
		(0.00)	(1.00)	(0.00)		(0.00)	(1.00)	(0.00)
	Panel C. Based on log-likelihood							
LogLik	-3803.77	-2793.37	-1640.50	-1642.03	-3911.22	-2811.41	-1627.97	-1630.75
	(0.00)	(0.00)	(0.82)	(0.18)	(0.00)	(0.00)	(1.00)	(0.00)
AIC	7617.54	5608.74	3315.00	3330.07	7832.43	5644.82	3289.94	3307.50
	(0.00)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)
BIC	7641.55	5661.57	3396.65	3440.52	7856.44	5697.65	3371.58	3417.96
	(0.00)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)
Ent	-	0.39	0.00	6.26	-	0.15	0.00	9.67
		(0.00)	(1.00)	(0.00)		(0.00)	(1.00)	(0.00)
BIC+Ent	-	5662.35	3396.65	3453.04	-	5697.95	3371.58	3437.31
		(0.00)	(1.00)	(0.00)		(0.00)	(1.00)	(0.00)
	Panel D. Model classification							
$CR(S^K, S^*)$	-	0.46	1.00	0.50	-	0.33	1.00	0.67
	Case 2				Case 3			
	Panel A. Poisson hierarchical approach for durations							
	Mean: 2.77		Median: 2.79		Mean: 2.79		Median: 2.79	
K	1	2	3	4	1	2	3	4
	Panel B. Based on Bayes factor							
$K-1/K$	-	2549.43	1034.37	-43.66	-	1767.29	1859.02	-41.72
		(0.00)	(1.00)	(0.00)		(0.00)	(1.00)	(0.00)
	Panel C. Based on log-likelihood							
LogLik	-3476.39	-2181.26	-1643.67	-1645.12	-4121.47	-3217.42	-2267.99	-2267.99
	(0.00)	(0.00)	(0.80)	(0.20)	(0.00)	(0.00)	(0.58)	(0.42)
AIC	6962.77	4384.52	3321.34	3336.25	8252.94	6456.83	4568.99	4581.97
	(0.00)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)
BIC	6986.78	4437.35	3402.34	3446.70	8276.95	6509.66	4650.63	4692.43
	(0.00)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)
Ent	-	0.00	0.17	8.25	-	0.46	0.00	9.65
		(1.00)	(0.00)	(0.00)		(0.00)	(0.99)	(0.01)
BIC+Ent	-	4437.35	3405.33	3463.20	-	6510.58	4650.63	4711.73
		(0.00)	(1.00)	(0.00)		(0.00)	(1.00)	(0.00)
	Panel D. Model classification							
$CR(S^K, S^*)$	-	0.67	1.00	0.34	-	0.49	1.00	0.49

Table 1. Continued

	Case 4				Case 5			
	Panel A. Poisson hierarchical approach for durations							
	Mean: 2.86		Median: 2.86		Mean: 2.68		Median: 2.67	
K	1	2	3	4	1	2	3	4
	Panel B. Based on Bayes factor							
$K-1/K$	-	3230.40	1335.89	-42.43	-	829.03	124.42	-27.12
		(0.00)	(1.00)	(0.00)		(0.00)	(1.00)	(0.00)
	Panel C. Based on log-likelihood							
LogLik	-3963.62	-2328.01	-1639.66	-1640.47	-619.63	-192.83	-118.34	-119.62
	(0.00)	(0.00)	(0.68)	(0.32)	(0.00)	(0.00)	(1.00)	(0.00)
AIC	7937.24	4678.02	3313.32	3326.93	1249.27	407.67	270.69	285.24
	(0.00)	(0.00)	(0.99)	(0.01)	(0.00)	(0.00)	(1.00)	(0.00)
BIC	7961.25	4730.85	3394.96	3437.39	1259.74	430.71	306.29	333.41
	(0.00)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)
Ent	-	0.00	0.00	5.77	-	0.04	0.00	0.00
		(0.00)	(1.00)	(0.00)		(0.00)	(1.00)	(0.00)
BIC+Ent	-	4730.85	3394.96	3448.93	-	430.79	306.29	333.41
		(0.00)	(1.00)	(0.00)		(0.00)	(1.00)	(0.00)
	Panel D. Model classification							
$CR(S^K, S^*)$	-	0.50	1.00	0.50	-	0.67	1.00	1.00

	Case 6a				Case 6b			
	Panel A. Poisson hierarchical approach for durations							
	Mean: 2.84		Median: 2.85		Mean: 2.85		Median: 2.87	
K	1	2	3	4	1	2	3	4
	Panel B. Based on Bayes factor							
$K-1/K$	-	1993.06	2184.75	-43.77	-	1962.19	2172.92	-43.92
		(0.00)	(1.00)	(0.00)		(0.00)	(1.00)	(0.00)
	Panel C. Based on log-likelihood							
K	1	2	3	4	1	2	3	4
LogLik	-3777.81	-2760.87	-1648.09	-1649.57	-4395.32	-3780.27	-2778.76	-1672.24
	(0.00)	(0.00)	(0.81)	(0.19)	(0.00)	(0.00)	(0.88)	(0.12)
AIC	75645.61	5543.74	3330.17	3345.13	7570.54	5579.53	3377.80	3390.48
	(0.00)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)
BIC	7589.62	5596.56	3411.81	3455.59	7594.55	5632.35	3459.44	3500.94
	(0.00)	(0.00)	(1.00)	(0.00)	(0.00)	(0.00)	(1.00)	(0.00)
Ent	-	0.22	0.00	7.41	-	0.29	0.00	7.52
		(0.00)	(1.00)	(0.00)		(0.00)	(0.99)	(0.01)
BIC+Ent	-	5597.01	3411.81	3470.41	-	5632.94	3459.44	3515.97
		(0.00)	(1.00)	(0.00)		(0.00)	(1.00)	(0.00)
	Panel D. Model classification							
$CR(S^K, S^*)$	-	0.63	1.00	0.46	-	0.57	1.00	0.32

Notes. Case 0 is the baseline scenario, Case 1 includes a long recession, Case 2 includes a short expansion, Case 3 refers to uncertain turning points, in Case 4 one cluster is generated with less specific turning points, in Case 5 all clusters are generated with a few number of specific turning points, Case 6a and Case 6b are generated from right-skewed and left-skewed distributions. For each case and each number of clusters K, Panel A reports the averages over the 1000 simulations of the mean and median number of clusters from Koop and Potter (2007). Panel B reports the average of twice the log of Bayes factor for $K-1$ versus K clusters, with $K = 2, 3, 4$. Panel C shows the marginal log likelihoods, the Bayesian AIC and BIC selection criteria, the entropy, and the BIC corrected for misclassification. For each model selection procedure, figures in brackets show the ratio of the number of times, given as a fraction of unity, that the model with K clusters is achieved over the 1000 simulations. Panel D reports the average classification rates, each of which measures the percentage of observations that are allocated into the true group (as determined by S^*) by the estimated allocation S^K.

Table 2: Model estimation performance

Measure	Case 0	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6a	Case 6b
Panel A. Estimation of means								
MSE_{μ}	2.17	1.50	2.17	4.33	2.00	1.50	10.17	9.00
Panel B. Estimation of variances								
MSE_{σ}	0.67	0.33	0.72	0.78	1.44	0.56	0.56	0.44
MSE_{trace}	0.67	0.67	0.67	1.67	6.00	0.34	0.33	0.67

Notes. See notes of Table 1. Mean Squared Errors (MSE) are computed as the averages over the 1000 simulations of the squared differences between the posterior means and the true parameters.

Table 3: Model selection

Panel A. Poisson hierarchical approach for durations						
Mean: 7.96			Median: 8			
Panel B. Bayes factors and model selection						
K	LogLik	AIC	BIC	Entropy	BIC-Entropy	Bayes factor $K-1$ / K
1	-530.40	1070.80	1080.84	-	-	-
2	-216.92	455.83	477.91	1.64	481.20	602.92
3	-179.98	393.95	428.08	0.98	430.03	49.84
4	-144.96	335.92	382.09	0.99	384.06	45.99
5	-117.94	293.87	352.09	0.66	353.41	30.00
6	-103.45	276.90	347.16	0.29	347.73	4.93
7	-77.30	236.61	318.91	0.00	318.91	28.25
8	-54.33	202.66	297.01	0.00	297.01	21.90
9	-53.48	212.96	319.35	0.00	319.35	-22.35

Notes. Panel A reports the mean and the median number of clusters from Koop and Potter (2007). In Panel B, the first column refers to the marginal log likelihoods; the second and third columns refer to the Bayesian AIC and BIC selection criteria; the third column shows the entropy; the fourth column shows the BIC corrected for misclassification; the last column reports twice the log of the Bayes factor for $K-1$ versus K clusters, with $K = 2, ..., 9$.

Table 4: Model estimation

NBER		MSMM		Deviation (in months)	
Peaks	Troughs	Peaks	Troughs	Peaks	Troughs
1960.04	1961.02	1960.02 (1959.04,1960.12)	1961.02 (1960.05,1962.03)	2	0
1969.12	1970.11	1969.11 (1969.03,1970.08)	1970.11 (1970.04,1971.07)	1	0
1973.11	1975.03	1974.02 (1973.09,1974.06)	1975.03 (1974.12,1975.06)	−3	0
1980.01	1980.07	1979.11 (1979.07,1980.04)	1980.06 (1980.03,1980.10)	2	1
1981.07	1982.11	1981.07 (1981.03,1981.11)	1982.10 (1982.07,1983.01)	0	1
1990.07	1991.03	1990.04 (1989.12,1990.08)	1991.01 (1990.08,1991.06)	3	2
2001.03	2001.11	2000.12 (2000.07,2001.05)	2002.03 (2001.08,2002.11)	3	−4
2007.12	2009.06	2007.10 (2007.02,2008.06)	2009.08 (2009.01,2010.03)	2	−2
sum				10	−2
average				1.25	−0.125

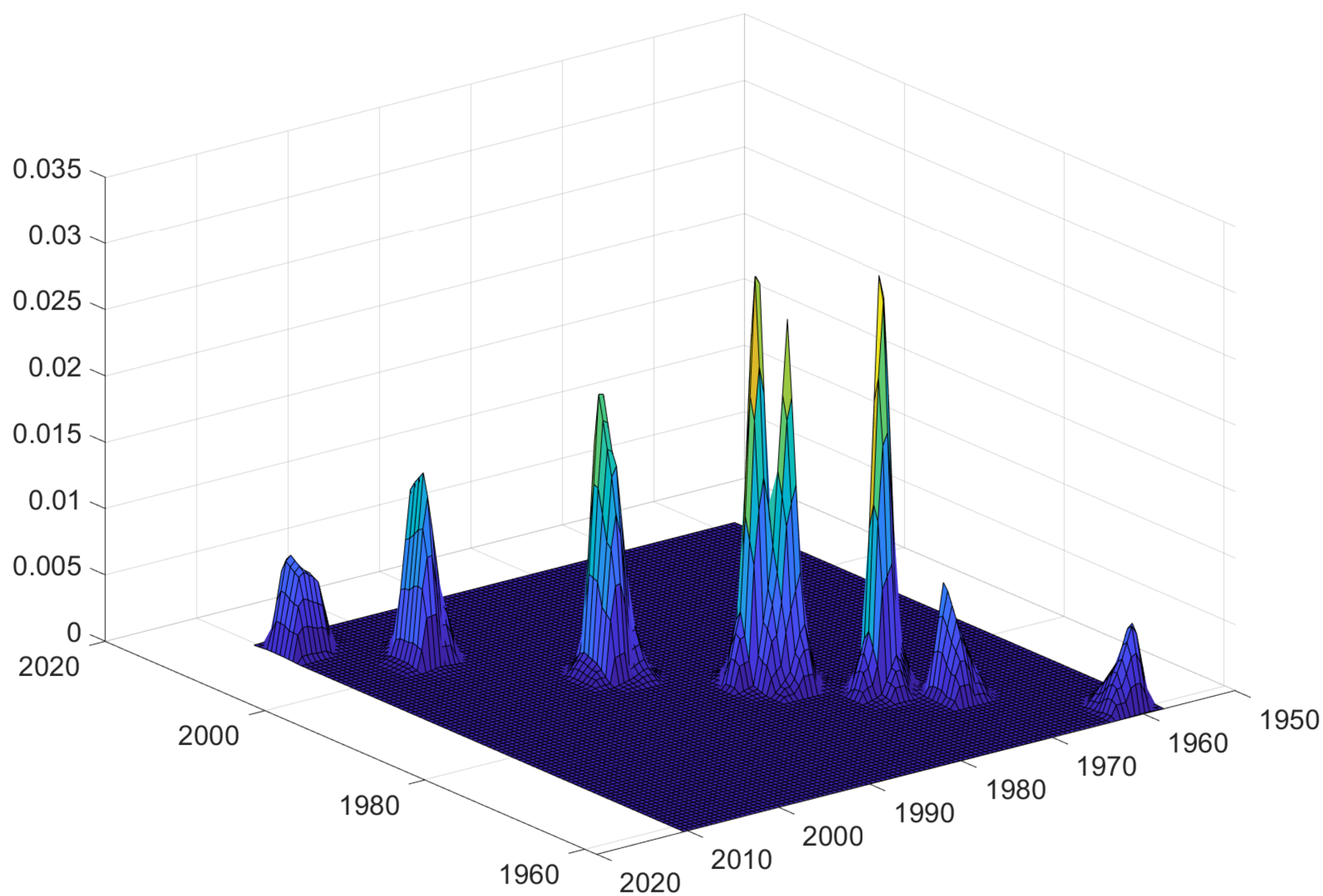
Notes. Columns 1 and 2 give the NBER-established turning point dates. Columns 3 and 4 report the turning point dates established with the mixture multiple change-point model, along with the range of values of the posterior probability distribution that includes 95% of the probability in brackets. Deviations are measured as the difference in months between the NBER turning points and those estimated by the model.

Table 5: Pseudo real-time analysis

NBER						MSMM					
Peak	Announce	Lag	Trough	Announce	Lag	Peak	Identify	Lag	Trough	Identify	Lag
1960.04	-	-	1961.02	-	-	1960.03 (1959.04,1961.01)	1966.01	-	1961.02 (1960.01,1962.01)	1966.01	-
1969.12	-	-	1970.11	-	-	1969.11 (1969.02,1970.07)	1971.06	18	1970.11 (1970.02,1971.12)	1971.06	7
1973.11	-	-	1975.03	-	-	1973.09 (1972.02,1974.11)	1974.08	9	1975.06 (1974.08,1976.12)	1975.11	8
1980.01	1980.06	5	1980.07	1981.07	12	1980.01 (1979.05,1980.08)	1980.07	6	1980.07 (1980.04,1981.05)	1981.01	6
1981.07	1982.01	6	1982.11	1983.07	20	1981.05 (1980.11,1981.10)	1982.03	8	1982.10 (1982.01,1983.09)	1983.05	6
1990.07	1991.04	9	1991.03	1992.12	21	1989.12 (1988.08,1991.04)	1990.03	-4	1990.11 (1990.01,1992.08)	1991.03	0
2001.03	2001.11	8	2001.11	2003.12	25	2000.12 (1999.10,2001.10)	2001.04	1	2001.11 (2000.10,2002.12)	2002.06	7
2007.12	2008.12	12	2009.06	2010.09	15	2007.10 (2005.10,2009.01)	2008.07	7	2009.06 (2008.07,2010.06)	2010.02	8

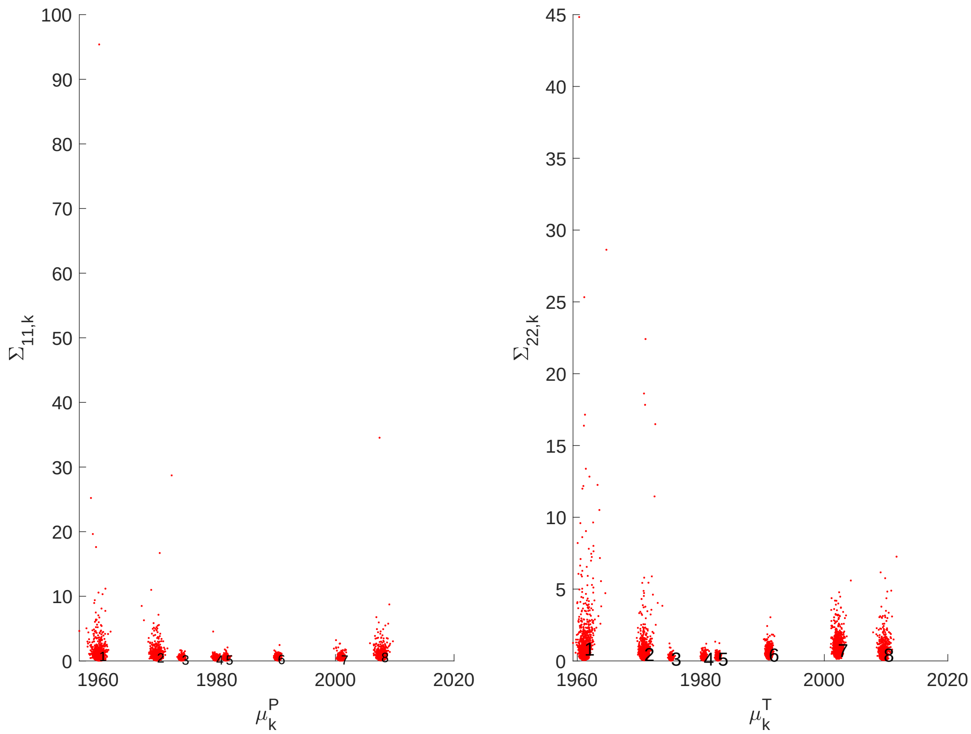
Notes. Columns labeled as ‘Announce’ refer to the date the NBER announced the turning points (prior to 1979, there were no formal announcements). Columns labeled as ‘Identify’ refer to the dates at which MSMM would have detected the turning points. Columns labeled as ‘Lag’ refer to the number of months after the NBER turning point date and the month that the NBER or MSMM announces the phase change. NBER-established turning points, means of the Gaussian distributions and their 95% credible intervals (in brackets) appear in columns labeled as ‘Peak’ and ‘Trough’.

Figure 1: Bivariate distribution of specific turning point dates

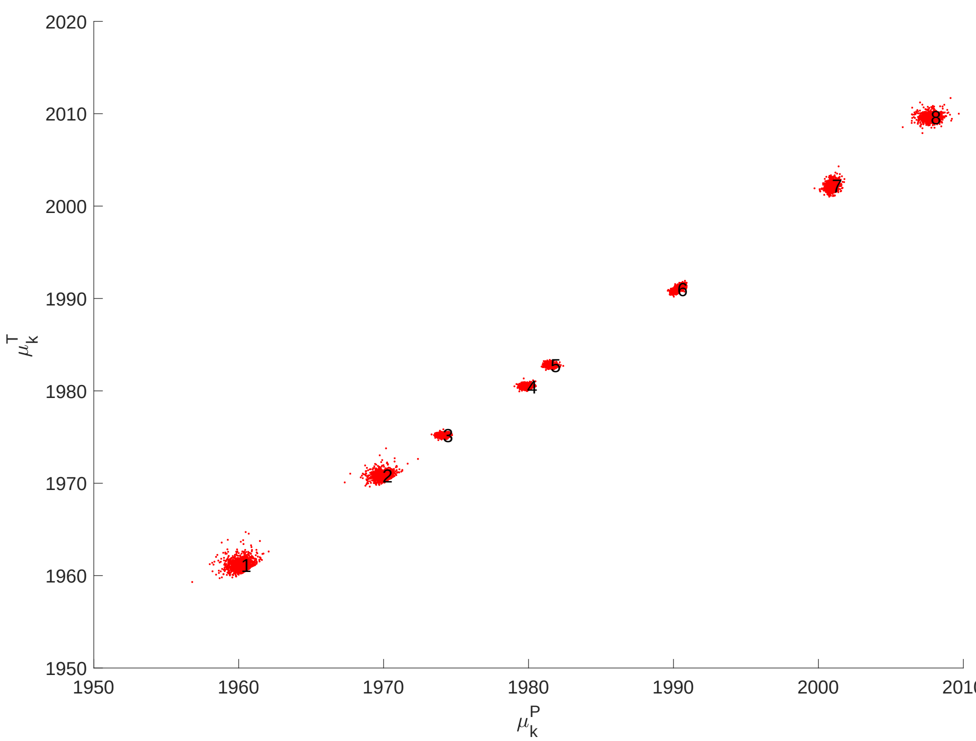


Notes. Kernel density estimate of the bivariate distribution of specific business cycle turning point dates.

Figure 2: Scatter plots of means and variances



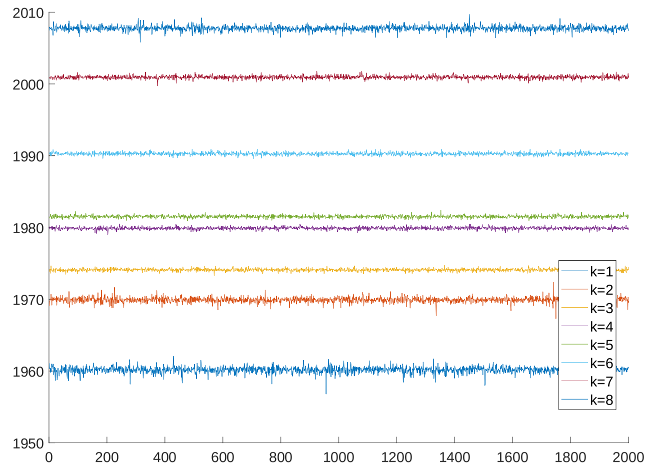
Panel A. Draws of $\left(\mu_k^P, \mu_k^T\right)$



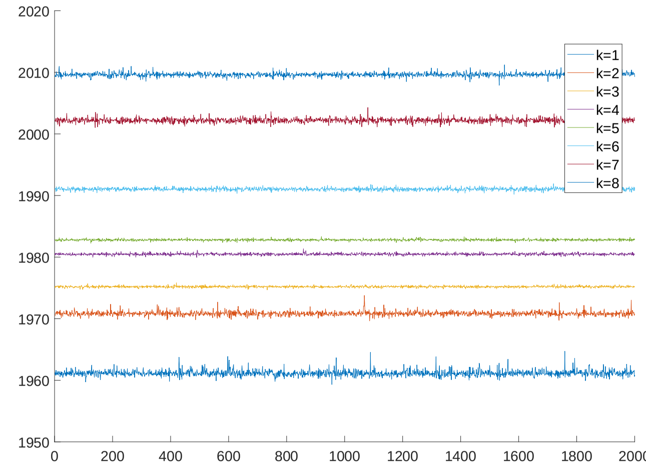
Panel B. Draws of $\left(\mu_k^P, \Sigma_{11,k}\right)$ and $\left(\mu_k^T, \Sigma_{22,k}\right)$

Notes. The figure displays the two-dimensional scatter plots of the MCMC draws for $\left(\mu_k^P, \mu_k^T\right)$, $\left(\mu_k^P, \Sigma_{11,k}\right)$ and $\left(\mu_k^T, \Sigma_{22,k}\right)$ for each of the $K = 8$ clusters.

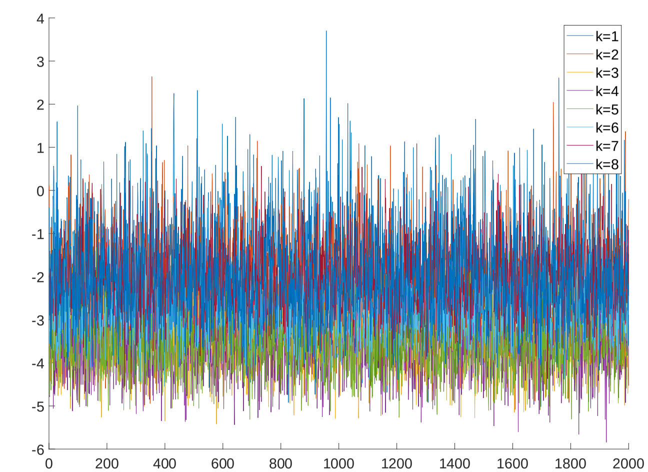
Figure 3: Behaviour of the Gibbs sampler



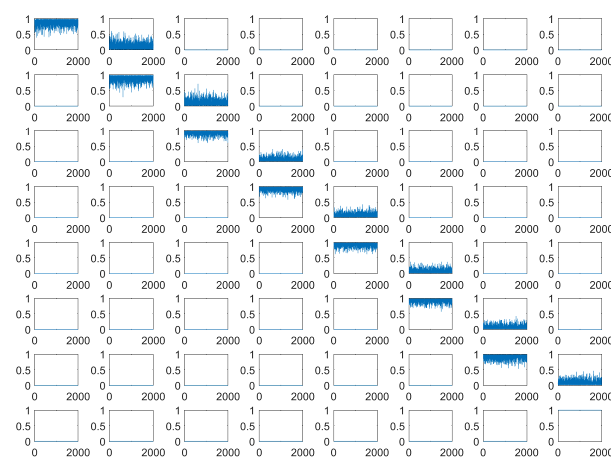
Panel A. Draws of μ_k^P



Panel B. Draws of μ_k^T



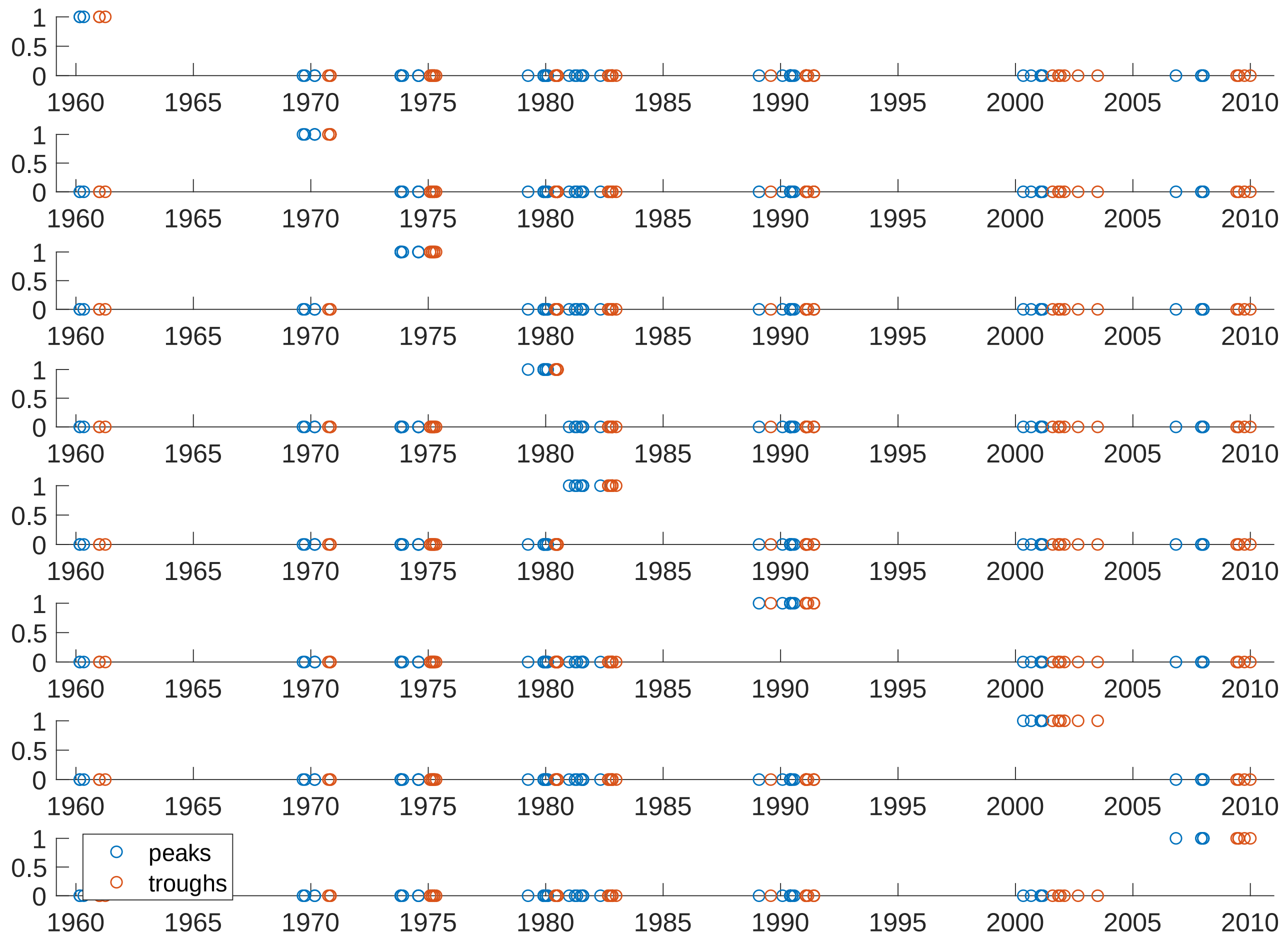
Panel C. Draws of $\log(|\Sigma_k|)$



Panel D. Draws of $p_{i,j}$

Notes. MCMC draws for μ_k^P, μ_k^T, $\log(|\Sigma_k|)$ and $p_{i,j}$ for each of the $K = 8$ clusters.

Figure 4: Classification probabilities



Notes. For each $k = 1, \ldots, 8$ and $i = 1, \ldots, N$, the figure plots the estimated posterior classification probabilities $\Pr(s_i = k|\theta)$ on a timeline that spans from January 1959 to August 2010. The probabilities represents the average number of times that observation i is allocated to cluster k across the M MCMC replications.