International Journal of Forecasting 37 (2021) 500–510

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Monitoring recessions: A Bayesian sequential quickest detection method[☆]

Haixi Li[a], Xuguang Simon Sheng[b,*], Jingyun Yang[c]

[a] Freddie Mac, United States of America
[b] American University, United States of America
[c] Rush University Medical Center, United States of America

ARTICLE INFO

Keywords:
Bayesian decision theoretic framework
Business cycle
Markov switching
Optimal stopping
Turning points

ABSTRACT

Monitoring business cycles faces two potentially conflicting objectives: accuracy and timeliness. To strike a balance between these dual objectives, we propose a Bayesian sequential quickest detection method to identify turning points in real time with a sequential stopping time as a solution. Using four monthly indexes of real economic activity in the United States, we evaluated the method's real-time ability to date the past five recessions. The proposed method identified similar turning-point dates as the National Bureau of Economic Research (NBER), with no false alarms, but on average, it dated peaks four months faster and troughs 10 months faster relative to the NBER announcement. The timeliness of our method is also notable compared to the dynamic factor Markov-switching model: the average lead time was about five months when dating peaks and two months when dating troughs.

1. Introduction

Economists in the United States have been pilloried for failing to predict the 2007–2009 recession. More recently, the Treasury and the Bank of England were embarrassed because both predicted an immediate post-Brexit vote recession that did not occur. These two examples and many others put the economic forecasting profession "to some degree in crisis" and illustrate the extreme difficulty in predicting or even identifying business-cycle turning points, e.g., Clements and Hendry (1999) and Giacomini and Rossi (2015).[1] The chronology of the United States business cycle is determined by the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER).[2] Due to the difficulty in determining whether a recession has started or ended, the NBER patiently waits

☆ This paper was presented at the 5th IAAE conference, St. Louis Fed conference on "*Central Bank Forecasting*," IIF workshop on "*Forecasting with Massive Data in Real Time*," Philadelphia Fed workshop on "*Early Warning Model*," 21st Dynamic Econometrics Conference, and Georgetown Center for Economic Research biennial conference. We thank Travis Berge, Marcelle Chauvet, Dean Croushore, Steven Durlauf, Neil Ericsson, Raffaella Giacomini, Domenico Giannone, David Hendry, George Monokroussos, Allan Timmermann, and the conference/workshop participants for very helpful comments. We also thank Michael McCracken (the editor), an associate editor, and two anonymous referees for their helpful comments and suggestions. Dr. Yang's research was supported by an NIH/NIA grant R01AG036042 and the Illinois Department of Public Health. The opinions expressed in this article are those of the authors and do not reflect the view of Freddie Mac.

* Correspondence to: 4400 Massachusetts Ave., N.W., Washington, DC 20016, United States of America.

E-mail address: sheng@american.edu (X.S. Sheng).

1 Clements and Hendry (1999) view structural breaks as the main source of forecast breakdowns. Giacomini and Rossi (2015) provide a comprehensive review on forecasting in non-stationary environments and illustrate what works and what does not in both reduced-form and structural models.

2 Burns and Mitchell (1946) laid the foundation for the concept of a business cycle, adopted by the NBER and other dating committees around the world. According to the NBER, "A recession is a period between a peak and a trough. During a recession, a significant decline in economic activity spreads across the economy and can last from a few months to more than a year".

https://doi.org/10.1016/j.ijforecast.2020.06.013
0169-2070/

for sufficient evidence to accumulate before making a decision. For example, the December 2007 and June 2009 business-cycle peak and trough were announced by the NBER after a year of due diligence. Exceptional accuracy comes at the cost of timeliness. Policymakers and business people are, however, interested in determining as quickly as possible whether the economy has entered or exited from a recession.

We frame monitoring business cycles as a Bayesian sequential quickest detection (Bsquid) problem to a two-state hidden Markov process. To derive the stopping time (i.e., the threshold), we specify a loss function that captures the dual requirements of timeliness and accuracy. Bayes' rule is used to update the probability of a regime switch. If the posterior probability of a regime switch exceeds the threshold, our Bsquid method identifies a turning point; otherwise, no break will be declared and the process will continue. The resulting threshold is state-dependent. When the null and alternative states differ considerably, the threshold tends to be higher, such that the decision maker can avoid false alarms without taking too much risk of a delay. By contrast, when two states are very close and it is hard to identify a break, the threshold becomes lower in order to avoid a delayed detection. This is one of the appealing features of the Bsquid method.

We applied the Bsquid method to dating business-cycle turning points, one of the most common breaks in economic time series. The current methodology, both in the academic literature and dating committees, is to date reference cycles using an aggregated statistic that summarizes the cyclical movement of the economy, e.g., De Mol et al. (2008) and Stock and Watson (2002). We followed this practice by constructing a common factor from four real-time monthly coincident indicators, namely, non-farm payroll employment, industrial production, real personal income excluding transfer receipts, and real manufacturing and trade sales. The Bsquid method identified the beginning of five recessions with reasonable accuracy, without "false alarms," and substantially faster than the NBER: the average lead time was about four months. Furthermore, our method showed systematic improvement over the NBER in the speed with which business-cycle troughs were identified. In particular, the Bsquid method announced the five troughs on average about 10 months ahead of the NBER announcement.

Our paper builds on the literature on quickest change detection. Earlier studies on this subject can be dated back to the 1930s (Shewhart, 1931). Wald (1947) proposed a sequential probability ratio test to reduce the number of sampling inspections without sacrificing the reliability of the final statistical decision. Since then, researchers have developed numerous methods to deal with similar problems in non-economic fields, e.g., Basseville and Nikiforov (1993), Lai (2001) and Poor and Hadjiliadis (2009). Chu et al. (1996) presented one of the early applications in economics for monitoring structural changes by introducing a fluctuation monitoring procedure based on recursive estimates of parameters and a cumulative sum procedure based on the behavior of recursive residuals. In this paper, we adopt the Bayesian approach instead, as it is a natural solution to process sequentially arrived information (West, 1986). Our study is most closely related to Shiryaev (1978), where the parameters in the distribution are assumed to be known and the univariate random sequence is assumed to be independent and identically distributed (iid). We generalize Shiryaev's work by allowing for non-iid univariate stochastic processes with unknown pre- and post-break parameters.

Our paper is also closely related to the literature on business-cycle turning points. Berge and Jorda (2011) and Stock and Watson (2014) focused on estimating turning points, conditional on a turning point having occurred. Our study differs from these papers in that we date turning points in real time, rather than establishing in-sample chronologies of business cycles. Many others have considered predicting turning points using leading economic and financial variables, e.g., Dueker (2005), Estrella and Mishkin (1998), Kauppi and Saikkonen (2008) and Rudebusch and Williams (2009). As pointed out by Hamilton (2011), this practice has not proved to be robust in its out-of-sample performance. The analysis in Giacomini and Rossi (2009) revealed the prime role played by instabilities in the data-generating process in causing forecast breakdowns. Berge (2015) went a step further and concluded that no model issued strong warning signals ahead of the 2001 and 2007 recessions.

By contrast, we pursue the modest goal of trying to identify a turning point soon after it has occurred. For example, Aastveit et al. (2016), Camacho et al. (2018), Chauvet and Hamilton (2006) and Chauvet and Piger (2008) advocated using a dynamic factor Markov-switching (DFMS) model to generate recession probabilities.[3] Compared to these papers, we explicitly model a decision maker's dual requirements of timeliness and accuracy and frame the problem of monitoring business cycles as a sequential stopping time. Our Bsquid framework is objective, transparent, and repeatable. The proposed method announced the past five recessions faster than the DFMS model: the average lead time was about five months when dating peaks and two months when dating troughs.

The rest of the paper is organized as follows. We describe the Bsquid method in Section 2. Section 3 introduces the data and discusses the empirical results of monitoring recessions. Section 4 concludes. Additional tables are relegated to the appendix.

2. Bayesian sequential quickest detection method

In this section, we start by describing the stochastic process of the underlying problem and then present the Bayesian decision theoretic framework for monitoring recessions.

For a univariate time series y_t, let $f(y_t|s_t = j, y_{t-1}, \theta)$ be its conditional density. We denote $\theta \in \Theta$ as a set of parameters for the parameterized distribution f, where Θ denotes the parameter space, s_t represents the state of

[3] Other algorithms in dating recessions include Giusto and Piger (2017) and Harding and Pagan (2006). Harding and Pagan (2006) proposed a non-parametric algorithm and Giusto and Piger (2017) introduced a simple machine-learning algorithm for dating recessions.

the economy, and j takes a value of 0 or 1, corresponding to the expansion and recession phase of the economy, respectively. Suppose that the economy starts at $s_0 = 0$ with probability π at $t = 0$, and changes to $s_\tau = 1$ at an unknown time τ. We assume a geometric prior distribution for the regime switching time τ:

$$P(\tau = k) = \begin{cases} \pi & \text{if } k = 0 \\ (1-\pi)\rho(1-\rho)^{k-1} & \text{if } k = 1, 2, \ldots \end{cases} \tag{1}$$

Let π_t denote the probability of a regime switch that has already occurred before t:

$$\pi_t = P(\tau \le t | \mathcal{F}_t, \theta), \tag{2}$$

where $\mathcal{F}_t$ is the information set available at time t. Given the geometric prior distribution, π_t evolves according to the following equation:

$$\pi_t = \frac{(\rho(1-\pi_{t-1}) + \pi_{t-1}) f(y_t | s_t = 1, \mathcal{F}_t, \theta)}{f(y_t | \mathcal{F}_t, \theta)}, \tag{3}$$

where

$$f(y_t|\mathcal{F}_t, \theta) = (1-\rho)(1-\pi_{t-1}) f(y_t|s_t = 0, \mathcal{F}_t, \theta) + (\pi_{t-1} + \rho(1-\pi_{t-1})) f(y_t|s_t = 1, \mathcal{F}_t, \theta).$$

The log-likelihood function $\mathcal{L}_k(\theta)$ given θ and data available at time k can be calculated as

$$\mathcal{L}_k(\theta) = \sum_{t=0}^{k} \log f\,(y_t|\mathcal{F}_t, \theta)\,. \tag{4}$$

We now present the Bsquid problem and propose a solution. Provided that a state will change at an unknown time τ, an agent's objective is to detect the change as soon as possible with the minimum risk of a false alarm. Formally, the agent's problem can be specified as

$$\inf_{T \in \mathcal{T}} E_{\Pi_0} \left\{ E_\theta \left\{ 1\,(T < \tau) + c\,(T - \tau)^+ \right\} \right\}. \tag{5}$$

The inner expectation E_θ is taken conditional on the set of parameters θ, including the parameter in the geometric prior ρ. The outer expectation E_{Π_0} is conditional on the prior over the parameter space Π_0. The first component $(E_{\Pi_0}\{E_\theta\{1\,(T < \tau)\}\})$ denotes the probability of a false alarm, and the second $(E_{\Pi_0}\left\{E_\theta\left\{c\,(T-\tau)^+\right\}\right\})$ is the expected length of delayed detection, multiplied by the controlling factor c. The parameter c reflects the agent's penalty on delayed detection relative to false alarms. A larger value of c increases the cost of a delayed detection.

The quickest change detection problem with an iid univariate random sequence was first studied by Shiryaev (1978). In his study, the parameters in the distribution θ, including ρ in the geometric prior of τ, are assumed to be known. These assumptions are too restrictive in many empirical applications. We generalize the method proposed by Shiryaev (1978) by allowing for unknown parameters of a pre- and post-change distribution and an unknown ρ in the geometric prior.

To our knowledge, there is no known solution to the optimization problem presented in Eq. (5). We fill this gap by proposing a sequential stopping time. Given the information set $\mathcal{F}_k$ at time k, let $p_0(\theta)$ denote the prior distribution of θ. Then, the posterior distribution over the parameter space Θ is

$$p_k(\theta|\mathcal{F}_k) \propto p_0(\theta)\mathcal{L}_k(\theta), \tag{6}$$

where the log-likelihood function $\mathcal{L}_k(\theta)$ is defined in Eq. (4). Let θ_k be the posterior mean as the estimate of the parameter θ:

$$\theta_k = E(\theta \mid \mathcal{F}_k). \tag{7}$$

Given θ_k, the optimization problem in Eq. (5) can be simplified as

$$\inf_{T_k \in \mathcal{T}} E_{\theta_k} \left\{ 1\,(T_k < \tau) + c\,(T_k - \tau)^+ \right\}. \tag{8}$$

To solve this problem, note that the objective function in Eq. (8) can be converted to a value function $v(\pi_k)$, where π_k is the posterior probability of a regime switch at time k[4]:

$$v(\pi_k) = \inf_{T_k \in \mathcal{T}} E_{\theta_k}(1 - \pi_{T_k} + c \sum_{i=k}^{T_k+k-1} \pi_i). \tag{9}$$

Furthermore, this value function can be written in a recursive form as

$$v(\pi_k) = \min\{1 - \pi_k, c\pi_k + E_{\theta_k} v(\pi_{k+1} \mid \pi_k)\}, \tag{10}$$

where π_{k+1}, given π_k, evolves according to Eq. (3). To solve the value function, we define Operator Q as

$$(Qv)(\pi_k) = \min\{1 - \pi_k, c\pi_k + E_{\theta_k} v(\pi_{k+1} \mid \pi_k)\},$$

and by iteration we obtain

$$v(\pi_k) = \lim_{n\to\infty} (Q^n v)(\pi_k). \tag{11}$$

The optimal stopping time can be defined as

$$T_k^* = \begin{cases} 1 & \text{if } \pi_k \ge \pi_k^* \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

where $v(\pi_k^*) = 1 - \pi_k^*$, and π_k^* is the optimal threshold that separates the state variable π_k into a continuation region $\{\pi_k : \pi_k < \pi_k^*\}$ and a stopping region $\{\pi_k : \pi_k \ge \pi_k^*\}$. The intuition is straightforward. If an immediate stopping results in less loss than the expected value of continuation, then it is time to stop; otherwise, it continues.[5] Fig. 1 illustrates the iteration of the value function and the optimal threshold. The optimal threshold, given information available at time k, evolves as a function of the updated parameter estimate θ_k. Accordingly, we define the sequential stopping time as

$$T^* = \inf\{k \ge 0 | T_k^* = 1\}. \tag{13}$$

3. Dating United States recessions in real time

This section starts with a description of the real-time dataset. Then, we discuss the results of identifying business-cycle turning points. Finally, we take a close look at a recent recession with an alternative overall economic activity index.

4 For details about this conversion, refer to Chapter 5 of Poor and Hadjiliadis (2009) or Chapter 4.3 of Shiryaev (1978).

5 See Shiryaev (1978) for the formal proof of the optimal stopping time with known parameters.

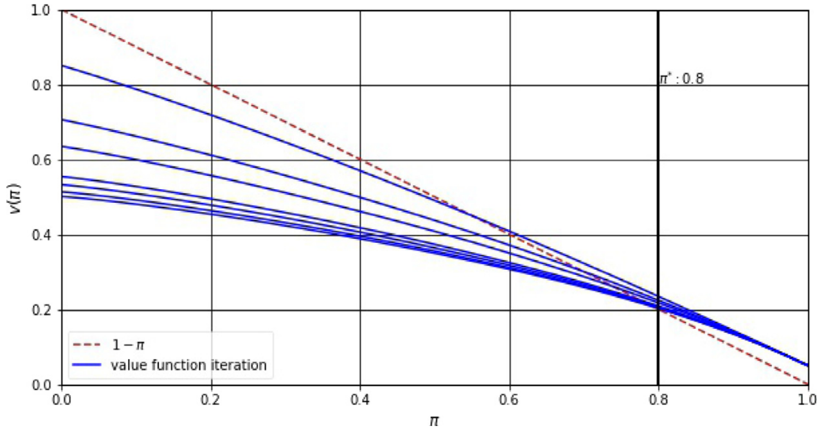

Fig. 1. The Value Function: Iteration and Convergence. Note: The value function $v(\pi)$ crosses the line $1-\pi$ at π^*, where π^* is the optimal threshold that separates the state variable π into a continuation region and a stopping region.

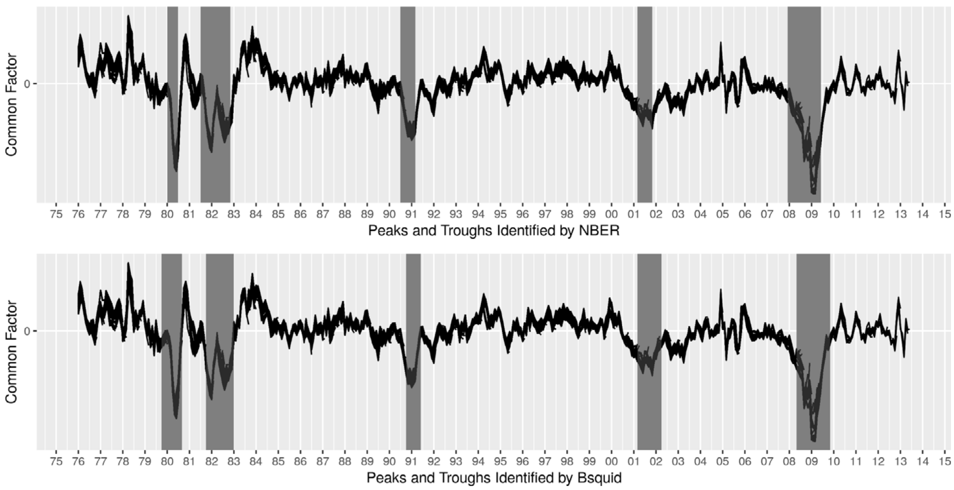

Fig. 2. Dating Recessions Using Four Monthly Coincident Indicators.

3.1. Data

We used four monthly coincident indicators: non-farm payroll employment (EMP), industrial production (IP), real personal income excluding transfer receipts (PIX), and real manufacturing and trade sales (MTS). The growth rates of these series are highlighted by the NBER in their decision on dating turning points. The dataset contains vintages of the four series from November 1976 to August 2013.[6] For each vintage, the sample period is from February 1967 until the most recent month available for that vintage. For the series of EMP, IP, and PIX, data are released with a one-month lag. For MTS, data are released with a two-month lag. To deal with this jagged data structure, we use Kalman filtering to fill in missing observations, and a dynamic factor model to extract a common factor.[7] This practice is consistent with the current literature by constructing an aggregated statistic that summarizes the cyclical movement of the economy; see, e.g., Stock and Watson (2014) and Hamilton (2011).[8]

Fig. 2 plots the common factor extracted from the four monthly coincident indicators using different data vintages. The shaded areas denote the recessions identified by the NBER (upper plot) and the Bsquid method (lower plot), which are discussed in greater detail in the next section.

6 Jeremy Piger kindly provides the real time dataset of these four series on his website. ALFRED can be used to extend the dataset past 2013.

7 In an earlier version of this paper, we filled in missing values using a smoothing spline; see, e.g. Shumway and Stoffer (2006). After obtaining the balanced data, we applied principal component analysis to extract a common factor from four coincident indicators. The resulting common factor was very similar to that extracted from a dynamic factor model, and thus omitted here.

8 Stock and Watson (2014) used the term "average then date" to describe this practice and Hamilton (2011) surveyed current literature on dating recessions using a single highly aggregated series such as GDP.

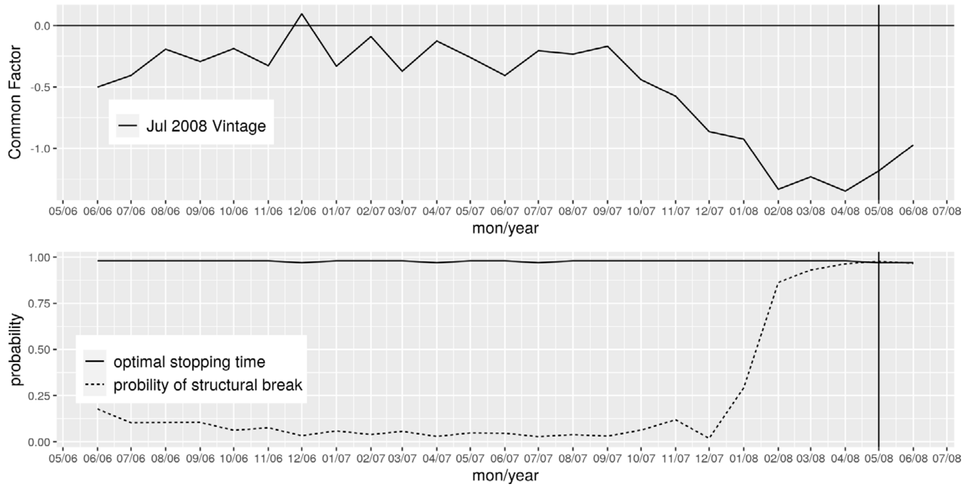

Fig. 3. Dating the Peak of May 2008 with Four Monthly Indicators.

3.2. Identifying turning points of business cycles

We applied the Bsquid method to date the past five recessions and compared the resulting turning points to those identified by the NBER. To this end, we focused on recessions that occurred after 1978, since the NBER made no formal announcements when it determined the dates of turning points before 1978.[9]

We used the Markov-switching model to characterize the cyclical movement of the economy:

$$\begin{aligned} & y_t = (1-\phi)\mu_{s_t} + \phi y_{t-1} + \epsilon_t, \\ & \epsilon_t \sim \mathcal{N}(0, \sigma^2), \\ & s_t \in \{0, 1\}, \end{aligned} \tag{14}$$

where y_t is the common factor extracted from many series; and $s_t \in \{0, 1\}$ denotes the state of the economy that changes at an unknown time $t = \tau$. Given the geometric prior and the sequence y_t, we calculated the probability of a regime switch that has occurred before t by Eq. (2). Based on the data available at time k, we obtained the log-likelihood function $\mathcal{L}_k(\theta)$ by Eq. (4), where the parameter set θ includes $\{\rho_0, \rho, \phi, \mu_0, \mu_1, \sigma\}$.

We used the levels of the common factor to date both peaks and troughs. Due to the lack of specific knowledge about the priors, we simply selected uniform distributions. We used historical data dating back to 1967 to identify the 1970 and 1973 recessions, as the purpose of this exercise was to calibrate phase-dependent prior distributions, as shown in Table 1.

Given the prior distributions and the likelihood function, we employed the Metropolis–Hastings algorithm to simulate the posterior distribution $p_k(\theta|\mathcal{F}_k)$. We then obtained the parameters θ_k^* as the posterior mean according to Eq. (7), and derived the sequential stopping time T_k^* as defined in Eq. (13). If the posterior probability of a regime switch exceeds the threshold (i.e. $\pi_k > T_k^*$), then the Bsquid method identifies a turning point; otherwise, no change is declared. Fig. 3 illustrates such a case when dating the peak of May 2008, where the probability of a regime switch is above the threshold based on the July 2008 data vintage.

To date peaks, we took the beginning of an expansion identified by the Bsquid method as given, and aimed at detecting the end of the expansion. Based on historical data, we assumed that the minimum length of an expansion is 12 months. Table 2 summarizes the results in dating peaks. The first column includes the peaks defined by the NBER. The second column includes the peaks identified by Chauvet and Piger (2008)'s DFMS model.[10] We obtained these dates directly from Chauvet and Piger (2008) for the first four recessions. For the 2007–2009 recession, we applied their rule to convert the recession probabilities into a zero/one variable that defines whether the economy is in an expansion or a recession regime. Comparing the NBER dates to those identified by our method in the third column illustrates the accuracy of the newly established dates. Out of five recessions, the Bsquid method identified the beginning of four with reasonable accuracy, within three months of the NBER date. Furthermore, our method produced no false-positive signals over the sample period. Column 4 shows the dates when the NBER announced that a peak had occurred, and Columns 5 and 6 show the corresponding dates announced by the DFMS model and our method, respectively. On average, the Bsquid method announced the peak faster than the NBER and the DFMS model. The average lead time for the five peaks in the sample was about four months ahead of the NBER and five months ahead of the DFMS model.

The 2007–2009 recession was an exception. The Bsquid method identified May 2008 as the peak, which was five months later than the NBER date. To understand this discrepancy, note that our method made this call using

9 The Business Cycle Dating Committee was created in 1978, and since then, there has been a formal process of announcing the NBER determination of a peak or trough in economic activity.

10 Using the same data, Chauvet and Piger (2008) showed that the real-time performance of the DFMS model outperformed those by Bry and Boschan (1971)'s nonparametric methodology in terms of both accuracy and timeliness. For this reason, we do not include the comparison to the Bry–Boschan method here.

Table 1
Prior distributions and control factors.

ρ	μ_0	μ_1	ϕ	σ	ρ_0	c
For dating peaks						
[0, 1]	[−0.20, 0.80]	[−2.00, −0.80]	[0.10, 0.40]	[0.50, 0.80]	0.05	0.01
For dating troughs						
[0, 1]	[−2.00, −1.00]	[−0.60, 0.80]	[0.10, 0.40]	[0.20, 0.60]	0.05	0.02

Table 2
Dating peaks in real time with four indicators.

Peaks identified by			Peaks announced by			Months ahead of	
NBER	DFMS	Bsquid	NBER	DFMS	Bsquid	NBER	DFMS
Jan 1980	Jan 1980	Oct 1979	Jun 1980	Jul 1980	Dec 1979	6 months	7 months
Jul 1981	Jul 1981	Oct 1981	Jan 1982	Feb 1982	Dec 1981	1 months	2 months
Jul 1990	Jul 1990	Oct 1990	Apr 1991	Feb 1991	Dec 1990	4 months	2 months
Mar 2001	Mar 2001	Mar 2001	Nov 2001	Jan 2002	May 2001	6 months	8 months
Dec 2007	Jan 2008	May 2008	Dec 2008	Jan 2009	Jul 2008	5 months	6 months

Table 3
Dating troughs in real time with four indicators.

Troughs identified by			Troughs announced by			Months ahead of	
NBER	DFMS	Bsquid	NBER	DFMS	Bsquid	NBER	DFMS
Jul 1980	Jun 1980	Sep 1980	Jul 1981	Dec 1980	Nov 1980	8 months	1 months
Nov 1982	Oct 1982	Jan 1983	Jul 1983	May 1983	Mar 1983	4 months	2 months
Mar 1991	Mar 1991	Jun 1991	Dec 1992	Sep 1991	Jul 1991	17 months	2 months
Nov 2001	Nov 2001	Apr 2002	Jul 2003	Aug 2002	May 2002	14 months	3 months
Jun 2009	Jul 2009	Nov 2009	Sep 2010	Jan 2010	Dec 2009	9 months	1 months

the July 2008 data vintage. The overall economy kept declining after June 2007 and became substantially negative in May 2008. Supporting evidence can also be found in the NBER announcement made in December 2008, "other series considered by the committee—including real personal income less transfer payments, real manufacturing and wholesale-retail trade sales, industrial production, and employment estimates based on the household survey—all reached peaks between November 2007 and June 2008". The Committee's decision was largely driven by payroll employment, which reached a peak in December 2007 and declined every month since then.[11]

To date troughs, we adopted the same structural model as above. We took the beginning of a recession identified by the Bsquid method as given, and aimed at detecting the end of the recession. Based on historical data, we assumed that the minimum length of a recession is six months. We present these results in Table 3. Our method identified five troughs relatively accurately, within five months of the NBER date. Most importantly, our method showed systematic improvement over the NBER in the terms of the speed with which these troughs were announced. On average, the Bsquid method announced the five business-cycle troughs 10 months ahead of the NBER announcement. The maximum lead time was 17 months for the 1991 trough.

The Bsquid method announced the trough faster than the DFMS model, and the average lead time for the five troughs was about two months. Despite a slight disadvantage in terms of timeliness, the DFMS model was accurate in identifying troughs, all within one month of the NBER date. To understand this notable accuracy, note that Chauvet and Piger (2008) adopted a conservative two-step approach to dating troughs. In the first step, they required that the probability of recession moves from above to below 20% and remains below 20% for three consecutive months before a new expansion phase is identified. In the second step, they identified the first month of this expansion phase as the first month prior to month t, for which the probability of recession moves below 50%. By contrast, the Bsquid method does not depend on an ad hoc rule to convert the recession probabilities into a zero/one decision. Based on the first data vintage, the Bsquid method identifies a trough whenever the probability of a regime switch first crosses the stopping time.

The results above were obtained with the controlling factor $c = 0.01$ for dating expansions, and $c = 0.02$ for dating recessions. These two values were calibrated using historical data to identify the 1970 and 1973 recessions.[12] The identified business-cycle dates varied little with alternative values of c, as shown in Tables A.1 and A.2. As expected, selecting a smaller value of c reduced the probability of false alarms, and a larger value led to earlier detections of recessions, albeit at the cost of accuracy.

To select the value for c in practice, note that $c \geq 0$ reflects the decision maker's penalty on delayed detection relative to false alarms. The probability of false alarms (i.e., the first term in Eq. (5)) takes a value from 0 to 1, and the expected length of delayed detection (i.e., the

11 For more details, refer to www.nber.org/cycles/dec2008.html.

12 If a decision maker changes the preference for accuracy versus timeliness when dating a particular recession, the value of c can be updated based on the information at the moment of the estimation. We leave this for future research.

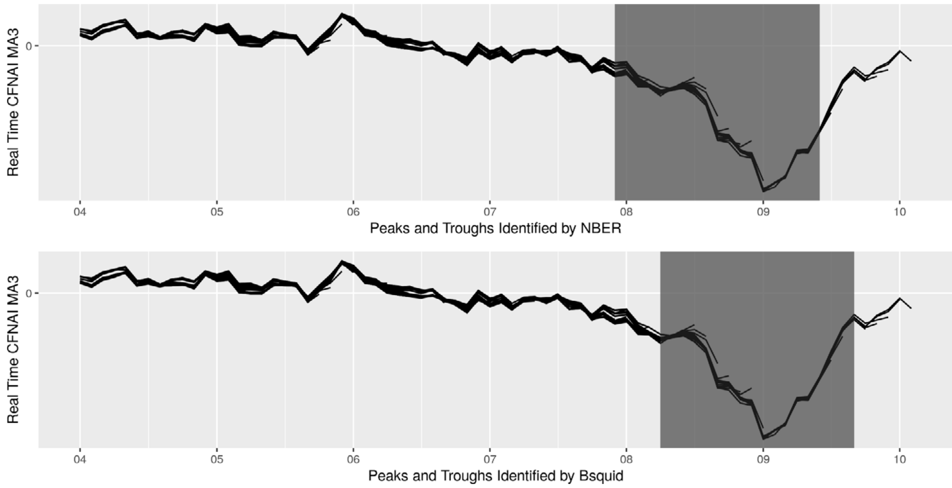

Fig. 4. Dating the 2007–2009 Recession Using Chicago Fed National Activity Index.

Table 4
Dating the 2007-2009 Recession in Real Time with CFNAI.

Identified by NBER	Identified by Bsquid	Leads (−) or lags (+)	Announced by NBER	Announced by Bsquid	Months ahead of NBER announce
Dec 2007	Apr 2008	+4 months	Dec 2008	May 2008	7 months
Jun 2009	Sep 2009	+3 month	Sep 2010	Oct 2009	11 months

second term in Eq. (5)) takes positive values that can be much larger than 1. To minimize both false alarms and the length of delayed detection multiplied by the controlling factor, the decision maker can choose a value of c that balances these two objectives. Our general suggestion is to estimate the model with a training sample and calibrate the value for the controlling factor.

To summarize, there are two main reasons why the turning points identified by the Bsquid method differ from those of the NBER. First, due to the difficulty in determining whether a recession has started or ended, the NBER patiently waits for sufficient evidence to accumulate before making a decision. This exceptional accuracy comes at the cost of timeliness. By contrast, the Bsquid method maintains a careful balance between these two conflicting objectives and identifies those turning points with reasonable accuracy, as soon as possible. Second, the data on four series are subject to serious revisions in real time, as aptly pointed out by Croushore and Stark (2001). The Bsquid method uses the first data vintage and makes the decision based on limited information available in real time. As a result, our method showed systematic improvement over the NBER in the speed with which business-cycle turning points were announced: the average lead time was about four months when dating peaking and 10 months when dating troughs.

3.3. Dating the great recession with an alternative index

So far, we applied the Bsquid method to four monthly series to date business-cycle turning points. It is possible to extract the information about national economic activity from a much larger set of monthly series. We explore this possibility here by using one monthly coincident indicator: the Chicago Fed National Activity Index (CFNAI). We selected this index because it provides accurate signals about the current state of the economy, as documented in Berge and Jorda (2011).

The CFNAI is a monthly index of United States economic activity constructed to summarize variations in 85 data series classified into four broad categories: production and income; unemployment and hours; personal consumption and housing; and sales, orders, and inventories. The index was designed as a coincident indicator of national economic activity and is an example of a "Goldilocks" index: the information from various series is combined to reflect deviations around a trend of economic growth. The index is normalized to have zero mean and unit standard deviation. When the value of the index is zero, this suggests that the economy is moving along a historical growth path. A negative value of the index is "cold" (growth is below average), and a positive value is "hot" (above average).[13]

We used the three-month moving average of this index to smooth the month-to-month variations over time in order to provide a more consistent picture of the cyclical movement of the economy. Moreover, since the CFNAI was not available in real time until January 2001, we used historical data from this index to calibrate the parameters. Specifically, we used the January 2001 data vintage to identify recessions that occurred in 1980, 1981, 1990, and 2001. For this exercise, the controlling factor was $c = 0.01$ for dating peaks, and $c = 0.03$ for dating troughs. Using these two values, we evaluated our method's performance at dating the 2007–2009 recession using the real-time index, as shown in Fig. 4.

Table 4 summarizes the results. The Bsquid method identified April 2008 as the beginning of the Great

13 The CFNAI index is publicly available at www.chicagofed.org/publications/cfnai/index.

Recession—about four months later than the NBER date, but one month earlier than the date based on the four monthly series (see Table 2). Furthermore, the Bsquid method announced the onset of the Great Recession in May 2008, faster than the announcement from the NBER in December 2008 and the four monthly series in July 2008; see Fig. B.1 for dating this peak. The Bsquid method identified September 2009 as the end of Great Recession—about three months later than the NBER date, but two months earlier than the date from the four monthly series (see Table 3); see Fig. B.2 for dating this trough. And our method showed substantial improvement over the NBER in terms of speed, with a leading time of about 11 months.

The superior performance of our method is mainly driven by the state-dependent threshold value, which is in a sharp contrast to using the fixed threshold of the CFNAI in identifying recessions. For example, the document posted on the Chicago Fed website suggests using −0.70 as the recession threshold in practice. Berge and Jorda (2011) found that an optimal threshold value of −0.72 maximizes the utility of the classification of this index into recessions and expansions by assuming equally weighted benefits of hits and costs of misses. Using either threshold value during the period from 1979–2011, the CFNAI signaled that the economy was in a recession within three months of the NBER date. Both thresholds, however, generated one false alarm in July 1989, when the index fell to −0.94 but no recession occurred. This brief analysis, despite using the most recently available data vintage, highlights the inherent problem of adopting a fixed threshold rule to identify recessions. The rule of thumb for this type considers only large deviations of the index from the mean of zero, but completely ignores other relevant information such as duration of these deviations. By contrast, our Bsquid method captures both the magnitude of the signal deviating from the null hypothesis (i.e., the overall amount of evidence against the null, or "strength") and the persistence of these deviations (i.e., the relative amount of evidence against the null, or "pattern").[14]

To further illustrate the difference between using a fixed threshold and state-dependent threshold, we performed Monte Carlo simulations. We generated a series y_t with Eq. (14). For all simulations, $\mu_1 = 0.5$, $\phi = 0.1$ and $\sigma = 0.3$. The only difference was in the parameter value for μ_0: $\mu_0 = -0.3$ under the first scenario, and $\mu_0 = -0.9$ under the second scenario. These two values describe situations where the null and the alternative states are quite similar (Scenario 1) or different from each other (scenario 2). At $t = 40$, we introduced a regime switch, such that y_t changed from the null to the alternative state. The Bsquid method correctly detected the change at $t = 41$ under both scenarios, as shown in Fig. 5. Notably, the threshold in Scenario 2 was on average higher than in Scenario 1. This is one of the desirable features of the Bsquid method: the threshold is state-dependent. When the null and alternative states are quite different (Scenario 2), the threshold is higher such that the decision maker avoids false alarms without taking too much risk of a delay. By contrast, when two states are very close and it is difficult to identify a regime switch (Scenario 1), the threshold is lower in order to avoid delayed detections.

4. Concluding remarks

We developed a Bayesian decision theoretic framework within which we clearly illustrated the decision maker's objectives of accuracy and timeliness. To achieve both objectives, we proposed a Bsquid method with a sequential stopping time. Monte Carlo simulations confirmed our method's ability to detect regime switches rapidly and without false alarms.

We presented two empirical examples of dating United States business-cycle turning points by using four monthly coincident indicators and the Chicago Fed National Activity Index. Our Bsquid method identified and announced business-cycle turning points at an impressive speed compared to the NBER announcement dates. In particular, our method announced the five peaks four months ahead of NBER and the five troughs about 10 months ahead. The timeliness of our method was also notable compared to the dynamic factor Markov-switching model: the average lead time was about five months when dating peaks and two months when dating troughs. Increased speed did not come with any sizable loss of accuracy.

Since our method applies to any economic and financial time series with regime switching, it has the potential to serve a wide variety of empirical applications, such as timely detection of financial stress and jumps in policy uncertainty. By applying the Bsquid method to empirical studies, practitioners would take a significant step toward real-time recognition of regime switching. Another worthwhile extension is to apply our method to monitor common regime changes in panel data settings. We leave this for future research.

Appendix A. Tables

See Tables A.1 and A.2.

Appendix B. Figures

See Figs. B.1 and B.2.

[14] It might be interesting to evaluate the performance of our method with approaches such as the receiver operating characteristic (ROC) curve. However, the ROC curve does not fit our scenario for two reasons. First, the ROC curve only assesses the performance of the posterior probability of a regime switch. It ignores the sequential optimal stopping time with which our method converts the posterior probability into a statistical decision. That said, our method generates a sequence of state-dependent thresholds. Consequently, our method only takes a single point on the ROC curve, corresponding to a true positive rate of 0.787 and a false positive rate of 0.047. Second, performance indices for a detection scheme include both accuracy and timeliness, where the latter is unique to literature on quickest change detection. The length of delay is a major performance index, and our method tackles this problem using a sequential stopping time, whereas the ROC curve fails to reflect this important property when detecting changes.

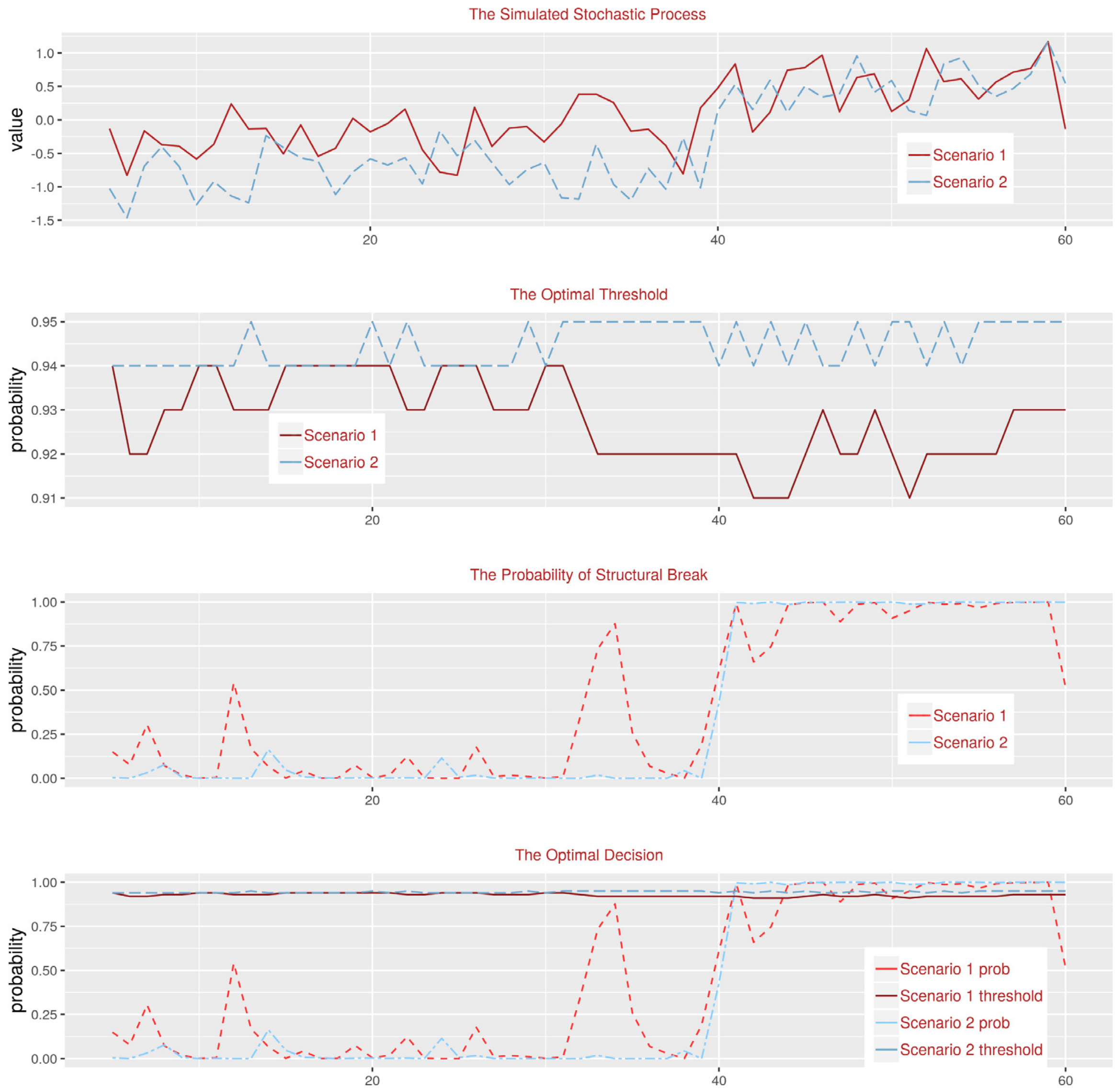

Fig. 5. Simulations under Scenario 1 ($\mu_1 = -0.3$) and Scenario 2 ($\mu_1 = -0.9$).

Table A.1
Dating peaks with four indicators using alternative control factors.

	$c = 0.02$		$c = 0.03$	
Identified by NBER	Identified by Bsquid	Announced by Bsquid	Identified by Bsquid	Announced by Bsquid
Jan 1980	Aug 1979	Sep 1979	Aug 1979	Sep 1979
Jul 1981	Oct 1981	Dec 1981	Oct 1981	Dec 1981
Jul 1990	Oct 1990	Nov 1990	Oct 1990	Nov 1990
Mar 2001	Feb 2001	May 2001	Dec 2000	Apr 2001
Dec 2007	Apr 2008	May 2008	Apr 2008	May 2008

Note: The benchmark for the control factor in dating peaks was $c = 0.01$. Using larger values led to earlier detection of recessions but at the cost of accuracy. Specifically, using either $c = 0.02$ or $c = 0.03$, the Bsquid method gave two false alarms: September 1992 and May 2003.

Table A.2
Dating troughs with four indicators using alternative control factors.

	$c = 0.01$		$c = 0.03$	
Identified by NBER	Identified by Bsquid	Announced by Bsquid	Identified by Bsquid	Announced by Bsquid
Jul 1980	Sep 1980	Nov 1980	Sep 1980	Oct 1980
Nov 1982	Feb 1983	Mar 1983	Jan 1983	Mar 1983
Mar 1991	Jun 1991	Jul 1991	Jun 1991	Jul 1991
Nov 2001	May 2002	Jun 2002	Mar 2002	Apr 2002
Jun 2009	Dec 2009	Jan 2010	Nov 2009	Dec 2009

Note: The benchmark for the control factor in dating troughs was $c = 0.02$. Using a larger value, such as $c = 0.03$, led to earlier detection of recessions but at the cost of one false alarm in May 1982.

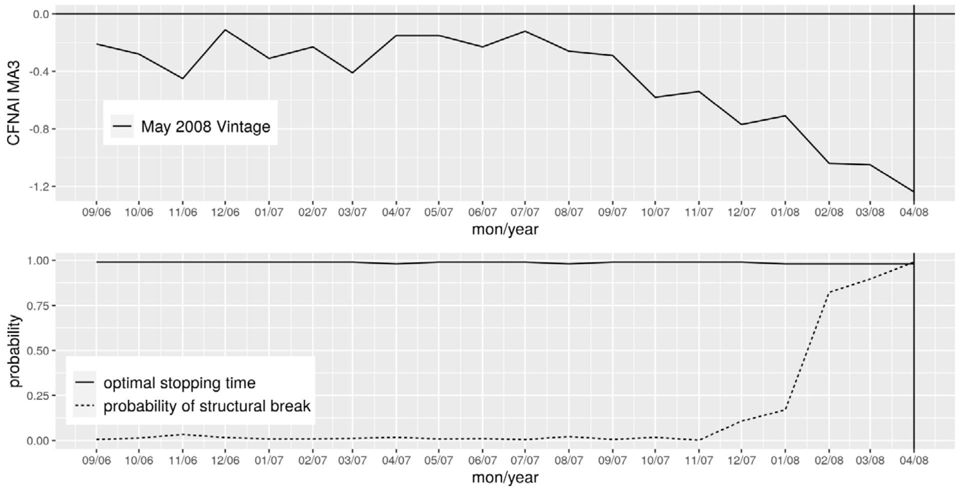

Fig. B.1. Dating the Peak of April 2008 with CFNAI.

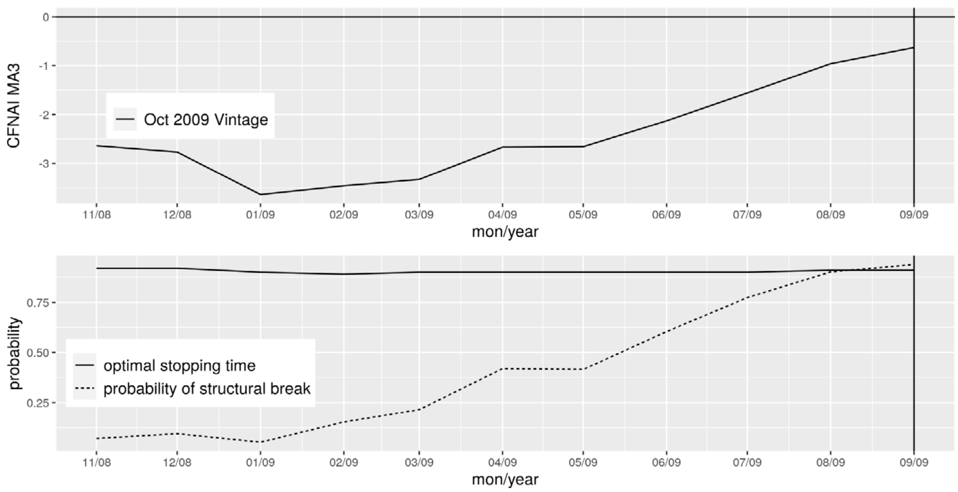

Fig. B.2. Dating the Trough of September 2009 with CFNAI.

References

Aastveit, K. A., Jore, A. S., & Ravazzolo, F. (2016). Markov-Switching dynamic factor models in real time. *International Journal of Forecasting*, *32*, 283–292.

Basseville, M., & Nikiforov, I. (1993). Detection of abrupt changes : theory and application. Englewood Cliffs, N.J: Prentice Hall.

Berge, T. (2015). Predicting recessions with leading indicators: Model averaging and selection over the business cycle. *Journal of Forecasting*, *34*, 455–471.

Berge, T., & Jorda, O. (2011). The classification of economic activity into expansions and recessions.. *American Economic Journal: Macroeconomics*, *3*, 246–277.

Bry, G., & Boschan, C. (1971). Cyclical analysis of time series: selected procedures and computer programs. In: NBER Technical Paper 20.

Burns, A. F., & Mitchell, W. C. (1946). Measuring business cycles. New York: NBER.

Camacho, M., Perez-Quiros, G., & Poncela, P. (2018). Markov-Switching dynamic factor models in real time. *International Journal of Forecasting*, *34*, 598–611.

Chauvet, M., & Hamilton, J. D. (2006). Dating business cycle turning points. In C. Milas, P. Rothman, D. van Dijk, & D. Wildasin (Eds.), *Contributions to economic analysis*: *vol. 276*, *Nonlinear time series analysis of business cycles* (pp. 1–54). Emerald Group, UK.

Chauvet, M., & Piger, J. (2008). A comparison of the real-time performance of business cycle dating methods. *Journal of Business & Economic Statistics*, *26*, 42–49.

Chu, C.-S. J., Stinchcombe, M., & White, H. (1996). Monitoring structural change. *Econometrica*, 1045–1065.

Clements, M., & Hendry, D. (1999). Forecasting non-stationary economic time series. Cambridge: The MIT Press.

Croushore, D., & Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, *105*, 111–130.

De Mol, C., Giannone, D., & Reichlin, L. (2008). Forecasting using a large number of predictors: Is bayesian shrinkage a valid alternative to principal components?. *Journal of Econometrics*, *146*, 318–328.

Dueker, M. (2005). Dynamic forecasts of qualitative variables: A qual var model of U.S. recessions. *Journal of Business & Economic Statistics*, *23*, 96–104.

Estrella, A., & Mishkin, F. (1998). Predicting U.S. recessions: Financial variables as leading indicators. *The Review of Economics and Statistics*, *80*, 45–61.

Giacomini, R., & Rossi, B. (2009). Detecting and predicting forecast breakdowns. *Review of Economic Studies*, *76*, 669–705.

Giacomini, R., & Rossi, B. (2015). Forecasting in nonstationary environments. *Annual Review of Economics*, *7*, 207–229.

Giusto, A., & Piger, J. (2017). Identifying business cycle turning points in real time with vector quantization. *International Journal of Forecasting*, *33*, 174–184.

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

Harding, D., & Pagan, A. (2006). Synchronization of cycles. *Journal of Econometrics*, *132*, 59–79.

Kauppi, H., & Saikkonen, P. (2008). Predicting U.S. recessions with dynamic binary response models. *The Review of Economics and Statistics*, *90*, 777–791.

Lai, T. L. (2001). Sequential analysis: some classical problems and new challenges. *Statistica Sinica*, *11*(2), 303–350.

Poor, V., & Hadjiliadis, O. (2009). Quickest detection. Cambridge University Press.

Rudebusch, G., & Williams, J. (2009). Forecasting recessions: The puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics*, *27*, 492–503.

Shewhart, W. A. (1931). Economic control of quality of manufactured product, vol. 27. (p. 501). America: D. Van Nostrand Company, Inc.

Shiryaev, A. N. (1978). Optimal stopping rules. Springer-Verlag.

Shumway, R. H., & Stoffer, D. S. (2006). Time series analysis and its applications with R examples. Springer-Verlag New York.

Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, *97*, 1167–1179.

Stock, J. H., & Watson, M. W. (2014). Estimating turning points using large data sets. *Journal of Econometrics*, *178*, 368–381.

Wald, A. (1947). Sequential analysis. John Wiley & Sons, New York..

West, M. (1986). Bayesian Model monitoring. *Journal of the Royal Statistical Society B*, *48*, 70–78.