Markov-switching dynamic factor models in real time*

Maximo Camacho
Universidad de Murcia
mcamacho@um.es

Gabriel Perez-Quiros
Banco de España and CEPR
gabriel.perez@bde.es

Pilar Poncela
European Commission, Joint Research Centre and Universidad Autónoma de Madrid
pilar.poncela@ec.europa.eu

Abstract

We extend the Markov-switching dynamic factor model to account for some of the specificities of the day-to-day monitoring of economic developments from macroeconomic indicators, such as mixed-sampling frequency and ragged-edge data. First, we evaluate the theoretical gains of using promptly available data to compute probabilities of recession in real time. Second, we show how to estimate the model that deals with unbalanced panels of data and mixed frequencies and examine the benefits of this extension through several Monte Carlo simulations. Finally, we assess its empirical reliability to compute real-time inferences of the US business cycle and compare it with the alternative method of forecasting the probabilities of recession from balanced panels.

Keywords: Business Cycles, Output Growth, Time Series.

JEL Classification: E32, C22, E27

*We are indebted to Marcelle Chauvet for kindly sharing part of the real-time data vintages used in the empirical application. We thank the editor, the associate editor and two anonymous reviewers for their comments. Part of this paper was written while the third author was visiting the Bank of Spain. Financial support from the Spanish government, contract grants ECO2015-70331-C2-1-R and ECO2016-76178-P (MINECO/FEDER), and 19884/GERM/15 (Groups of Excellence, Fundación Séneca, and Science and Technology Agency), is gratefully acknowledged. Any errors are our responsibility. The views in this paper are those of the authors and do not represent the views of the European Commission, the Bank of Spain or the Eurosystem. Codes and data that replicate the results can be downloaded from the authors' websites.

1 Introduction

The 2008-2009 was the most sustained economic slump that the United States has weathered since World War II. One of the lessons that this Great Recession left for economists was that policymakers and business people, who had become accustomed to the serene conditions of the Great Moderation, have dramatically increased their interest in determining as quickly as possible whether the economy has suffered from a business cycle phase shift. In this context, time-series models, which are able to automate the increasing complexity of the signal extraction problem in economics, help the economic agents to perform and update their real-time views of the developments in economic activity. These models deal with economic indicators that share the two properties of the business cycle documented early by Burns and Mitchell (1946): their signals about economic developments are spread over the different aggregates and they exhibit business cycle asymmetries.

Diebold and Rudebusch (1996) were the first to suggest a unified model that captures these two business cycle features from a set of economic indicators. They argued that comovements among the individual economic indicators can be modelled by using the linear coincident indicator approach described in Stock and Watson (1991), while the existence of two separate business cycle regimes can be modelled by using the Markov-switching specification advocated by Hamilton (1989). Integrating these approaches, Kim and Yoo (1995), Chauvet (1998) and Kim and Nelson (1998) combined the dynamic-factor and Markov-switching frameworks to propose different versions of statistical models which simultaneously capture both comovements and regime shifts. Camacho, Perez-Quiros and Poncela (2015) find that the fully non-linear multivariate specification outperforms the "shortcut" of using a linear factor model to obtain a coincident indicator which is then used to compute the Markov-switching probabilities. Recently, Chauvet and Hamilton (2006), Chauvet and Piger (2008), and Hamilton (2011) have examined the empirical reliability of these models in computing real-time inferences of the US business cycle states.

An important limitation of these Markov-switching dynamic factor models (MS-DFM) is that they were originally designed to deal with balanced panels of business cycle indicators. This crucial assumption means MS-DFM exhibit two drawbacks when applied to the (timely) day-to-day monitoring of economic activity in real time. The first drawback is that the typical lack of synchronicity in the flow of macroeconomic information implies that some indicators are published with a time delay, which requires dealing with unbalanced panels of data. Not accounting for this publication pattern would imply that the users of traditional MS-DFM who develop early assessments of economic developments from balanced panels of data will unavoidably incur one of the two following substantial costs. The first cost appears when the forecasts are made from the latest available balanced panel. In this case, the forecasts lose the latest and most valuable information contained in the promptly issued indicators at the time of the assessments. The second cost is that of being late when the analysts decide to wait until all the business cycle indicators become available and the inferences, which are then actually referred to the past. A significant example of this limitation is the two-month lag in the reporting real-time recession probability chart released by the St Louis Fed, which uses the MS-DFM originally developed in Chauvet (1998).

The second drawback of standard MS-DFM is that they relate variables sampled at the same frequency. In practice, although some of the macroeconomic indicators that are observed to infer business cycle states are sampled quarterly, some others, potentially useful in real-time inferences, are often sampled at a higher frequency. For example, the National Bureau of Economic Research (NBER) Dating Committee acknowledges that recessions are defined as significant declines in economic activity normally visible in real Gross Domestic Product (GDP), which is available quarterly, and real income, employment, industrial production, and wholesale-retail sales, which are available monthly.

In this paper, we examine the extent to which the incoming information provided by new releases of promptly published economic indicators, which potentially sampled at different frequencies, could help to improve the real-time inference about the business cycle. Using a theoretical MS-DFM, we show the extent to which inferences about the state of the economy can be improved upon by including early available indicators. Interestingly, we find that the improvements in performance depend on the factor loadings, the idiosyncratic variances, the dynamics of the common factor, and the difference between the means in the business cycle states.

Then, we extend the MS-DFM to allow economic agents that track business cycle developments in real time to use whatever business cycle economic indicator, regardless of their publication delays, and regardless of whether they are sampled monthly or quarterly. Based on a Markov-switching extension of the linear dynamic factor model proposed by Mariano and Murasawa (2003), our procedure deals with missing observations by using a time-varying nonlinear Kalman filter. Whenever the data is not observed, the missing observations are replaced by random draws from a variable whose distribution cannot depend on the parameter space that characterizes the Kalman filter. The corresponding row is then skipped in the Kalman recursion and the measurement equation for the missing observation is set to the random choice.

By means of several Monte Carlo experiments, we measure the magnitude of the gains of using our extension to compute business cycle inferences. We show that our proposal outperforms MS-DFM that requires balanced panels, especially when the forecasting horizon increases, when the two states are well separated, and when the variance and the inertia of the idiosyncratic components are low. Finally, we use a real-time data set to show that our extension of the MS-DFM leads to significant improvements in computing real-time business cycle inferences compared with forecasting from balanced and/or lagged panels of indicators using the four constituent monthly series of the Stock-Watson coincident index. Notably, we show that adding GDP does not seem to produce significant improvements in this setting.

The structure of this paper is organized as follows. Section 2 assesses the real-time features of the dataflow within a factor model framework. Section 3 examines the relative performance gains of dealing with ragged-edge data through a Monte Carlo experiment. Section 4 illustrates these results for US real-time data. Section 5 concludes.

2 The model

2.1 Model features

Our framework is the single-index Markov-switching dynamic factor model proposed in the mid-nineties by Kim and Yoo (1995), Chauvet (1998), and Kim and Nelson (1998), which incorporates both comovements and business-cycle shifts into a statistical model. The model postulates that a vector of N economic indicators, $\mathbf{y}_t = (y_{1,t}, ..., y_{N,t})'$, which are hypothesized to move contemporaneously with the overall economic conditions, can be decomposed as the sum of two components. The first component is a linear combination of r unobserved factors, $\boldsymbol{f}_t = (f_{1,t}, ..., f_{r,t})'$, that accounts for the common comovements. The second component is the $N \times 1$ time series vector $\mathbf{u}_t$, that represents the idiosyncratic movements in the series. This suggests the formulation:

$$\mathbf{y}_t \quad = \quad \boldsymbol{\Lambda} \quad \boldsymbol{f}_t \quad + \quad \mathbf{u}_t, \tag{1}$$

where $\boldsymbol{\Lambda}$ is the $N \times r$ factor loading matrix and $\mathbf{u}_t$ is the vector of idiosyncratic components.

To account for the business cycle asymmetries, it is assumed that the dynamic behavior of the common factors is governed by an unobserved regime-switching state variable, s_t. Within this framework, one can label $s_t = 0$ and $s_t = 1$ as the expansion and recession states at time t. In addition, it is standard to assume that the state variable evolves according to an irreducible 2-state Markov chain whose transition probabilities are defined by

$$p(s_t = j | s_{t-1} = i, s_{t-2} = h, ..., I_{t-1}) = p(s_t = j | s_{t-1} = i) = p_{ij}, \tag{2}$$

where $i, j = 0, 1$, and I_t is the information set up to period t, i.e., $\{\mathbf{y}_1, ..., \mathbf{y}_t\}$. We assume that the vector of common factor follows an autoregressive process with switching intercept,

$$\boldsymbol{f}_t = \boldsymbol{\mu}_{s_t} + \boldsymbol{\phi}_1 \boldsymbol{f}_{t-1} + ... + \boldsymbol{\phi_p} \boldsymbol{f}_{t-p} + \boldsymbol{a}_t, \tag{3}$$

where $\boldsymbol{a}_t$ is white-noise with variance Σ_a, which is independent of s_t. The main identifying assumption in the model expresses the core notion that the comovements of the multiple time series arise from the common component. This is achieved by assuming that $\mathbf{u}_t$ and $\boldsymbol{f}_t$ are mutually uncorrelated at all leads and lags. [1]

Each element $u_{i,t}$ of the idiosyncratic error $\mathbf{u}_t = (u_{1,t}, ..., u_{N,t})'$ is assume to follow an autoregressive process of order p_i

$$u_{i,t} = \psi_{i,1} u_{i,t-1} + ... + \psi_{i,p_i} u_{i,t-p_i} + \epsilon_{i,t}, \tag{4}$$

[1] Additional identifying restrictions, which are required to estimate the model, are not discussed in this theoretical section.

being $\{\epsilon_{i,t}\}$ a white noise process with variance σ^2_{i,s_t}. In matrix form, the dynamics of the idiosyncratic component is

$$\mathbf{u}_t = \mathbf{\Psi}_1 \mathbf{u}_t + ... + \mathbf{\Psi}_P \mathbf{u}_{t-P} + \boldsymbol{\epsilon}_t, \quad (5)$$

where $\mathbf{\Psi}_j = diag(\psi_{1,j}, ..., \psi_{N,j})$, $P = \max(p_1, ..., p_N)$, and $var(\boldsymbol{\epsilon}_t) = diag(\sigma^2_{1,s_t}, ..., \sigma^2_{N,s_t})$.[2]

2.2 Theoretical gains of using promptly available data

To facilitate the theoretical analysis, in this section we consider several simplifying assumptions. The first one is that we rely on publication lags of only one month. Therefore, we focus on considering the gains of using new partial information arriving at $t+1$. The second simplifying assumption is that the vector of idiosyncratic components, $\mathbf{u}_t$, is a multivariate white noise with variance-covariance matrix $\mathbf{\Sigma}_{u,s_t}$. However, we do maintain the assumption that this variance-covariance matrix can depend on the state. Third, the maximum lag length for the autoregressive model of the factors is $p = 1$.[3]

Then, we focus on computing inferences about the business cycle regime at $t+1$, that is, computing $prob(s_{t+1} = j)$ conditional on different information sets. Let us assume that all the indicators are collected up to time t, $\mathbf{y}_t$. However, we assume that only a subset of promptly published indicators, collected in the subvector $\mathbf{y}_{k,t+1}$, with $k < N$, is available at $t+1$. If we are restricted to using balanced panels, as in traditional MS-DFM, this forced us to use only the information I_t, obtained before the first arrival of new information at $t+1$. Therefore, our guess of the probability of being in a certain state at $t+1$ is the one step ahead forecast of the filtered probability $prob(s_t = j|I_t)$:

$$prob(s_{t+1} = j|I_t) = \sum_{i=0}^{1} prob(s_t = i|I_t)p_{ij}. \quad (6)$$

However, the extension of MS-DFM proposed in this paper allows for incorporating the information provided by $\mathbf{y}_{k,t+1}$ as it arrives. In this case, the inference of being in a

[2] Assuming that the idiosyncratic components are uncorrelated in cross-section is a necessary condition to identify the common factors in small scale factor models (Barhoumi et al., 2014). In large scale models ($N \to \infty$), the assumption is that idiosyncratic errors are only weakly correlated.

[3] The extensions to larger reporting lags, serially correlated idiosyncratic components, and further lags in the autoregressive model of the common factor are conceptually easy, although they would complicate computations considerably.

certain state at $t+1$ can be obtained as

$$prob(s_{t+1}=j|I_t,\mathbf{y}_{k,t+1})=\frac{f(\mathbf{y}_{k,t+1}|s_{t+1}=j,I_t)}{f(\mathbf{y}_{k,t+1}|I_t)}prob(s_{t+1}=j|I_t). \quad (7)$$

Using the new information will be useful if it raises the ability to increase the true signals. This implies that it should increase the probability of a given state when the economy is actually in that state. For instance, let us assume that the economy is in recession at $t+1$, i.e., $s_{t+1}=1$,[4] and that $0<prob(s_{t+1}=j|I_t)<1$.[5] Hence, from (7), the partial information provided by $\mathbf{y}_{k,t+1}$ is helpful to reduce false signals when $prob(s_{t+1}=1|I_t,\mathbf{y}_{k,t+1})>prob(s_{t+1}=1|I_t)$, which occurs whenever

$$f(\mathbf{y}_{k,t+1}|s_{t+1}=1,I_t)>f(\mathbf{y}_{k,t+1}|I_t). \quad (8)$$

Using the total law of probabilities, if $s_{t+1}=1$ the condition in (8) would be

$$f(\mathbf{y}_{k,t+1}|s_{t+1}=1,I_t)>f(\mathbf{y}_{k,t+1}|s_{t+1}=1,I_t)prob(s_{t+1}=1|I_t)+f(\mathbf{y}_{k,t+1}|s_{t+1}=0,I_t)prob(s_{t+1}=0|I_t), \quad (9)$$

which, rearranging terms, leads to

$$f(\mathbf{y}_{k,t+1}|s_{t+1}=1,I_t)>f(\mathbf{y}_{k,t+1}|s_{t+1}=0,I_t). \quad (10)$$

In practice, these two density functions may overlap and the condition does not hold for all possible values of $\mathbf{y}_{k,t+1}$. In these cases, even when the true state is $s_{t+1}=1$, it might happen that for some (relatively high) values of $\mathbf{y}_{k,t+1}$, $f(\mathbf{y}_{k,t+1}|s_{t+1}=1,I_t)<f(\mathbf{y}_{k,t+1}|s_{t+1}=0,I_t)$. This might lead to a false signal detection since the probability for $s_{t+1}=1$ decreases when $\mathbf{y}_{k,t+1}$ is observed. Accordingly, the usefulness of $\mathbf{y}_{k,t+1}$ to compute business cycle inferences should be evaluated on average. Taking natural logarithms, the condition in (10) implies that using the incoming information helps in inferring recession probabilities in actual recessions when

$$\ln f(\mathbf{y}_{k,t+1}|s_{t+1}=1,I_t)-\ln f(\mathbf{y}_{k,t+1}|s_{t+1}=0,I_t)>0. \quad (11)$$

[4] If we assumed that the true state is $s_{t+1}=0$, we would obtain symmetric results for the probability of expansion.

[5] It is straightforward to check that if $prob(s_{t+1}=j|I_t)=1$, then $prob(s_{t+1}=j|I_t,\mathbf{y}_{k,t+1})=1$, $j=0,1$.

Now, taking into account all possible outcomes of $\mathbf{y}_{k,t+1}$ for the true state $s_{t+1} = 1$, the expected value of the difference between the two conditional densities under conditional Gaussianity is given by the Kullback-Leibler divergence

$$\int \ln \frac{f(\mathbf{y}_{k,t+1}|s_{t+1}=1, I_t)}{f(\mathbf{y}_{k,t+1}|s_{t+1}=0, I_t)} f(\mathbf{y}_{k,t+1}|s_t = 1, I_t) d\mathbf{y}_{k,t+1}, \tag{12}$$

where we use the notation in Cover and Thomas (2006) for the multiple integral. The next proposition, which is based on the concept of conditional entropy, quantifies the potential advantage of adding the advanced business cycle signals provided by the promptly published indicators $\mathbf{y}_{k,t+1}$.

Proposition 1 *Assume the factor model given by (1), (2) and (3), where the vector of idiosyncratic components,* $\mathbf{u}_t$ *is multivariate white noise, and the lag length of the factors dynamics is* $p = 1$*, with autoregressive matrix* $\boldsymbol{\phi}$*. Under conditional Gaussianity, the gain in* $t+1$ *from observing the subvector of* k *variables* $\mathbf{y}_{k,t+1}$*is given by the Kullback-Leibler divergence of* $f(\mathbf{y}_{k,t+1}|s_{t+1} = 1, I_t)$ *with respect to* $f(\mathbf{y}_{k,t+1}|s_{t+1} = 0, I_t)$

$$KL = \frac{1}{2}\ln(2\pi)^k\left(|\Sigma_0| - |\Sigma_1|\right) + \frac{1}{2}tr\left(\Sigma_1\Sigma_0^{-1}\right) - \frac{k}{2} + \frac{1}{2}\left(\boldsymbol{f}_1 - \boldsymbol{f}_0\right)'\Lambda_k'\Sigma_0^{-1}\Lambda_k\left(\boldsymbol{f}_1 - \boldsymbol{f}_0\right), \tag{13}$$

where $\boldsymbol{f}_i = E\left(\boldsymbol{f}_{t+1}|s_{t+1} = i, I_t\right)$*,* $\boldsymbol{\Sigma}_i = E\left((\boldsymbol{y}_{k,t+1} - \Lambda_k\boldsymbol{f}_i)(\boldsymbol{y}_{k,t+1} - \Lambda_k\boldsymbol{f}_i)'|s_{t+1} = i, I_t\right)$*,* $i = 0, 1$ *and* Λ_k *is the* $k \times r$ *submatrix of factor loadings associated to* $\boldsymbol{y}_{k,t+1}$*.*

Proof: The proof is given in the Appendix.

This expression suggests that using the first arrival of new information collected in $\mathbf{y}_{k,t+1}$ is expected to raise the ability to increase the true business cycle signals when the difference in the expected value of the common factors and the between the within-state covariance matrices of the early available indicators are large.

As, for example, in Chauvet and Hamilton (2006), Chauvet and Piger (2008), and Hamilton (2011), empirical applications of MS-DFM typically rely on the assumptions that there is only one common factor and that the variances are state-independent. The next corollary focuses in this case.

Corollary 2 *For the one-factor model with state-independent variances,* $\boldsymbol{\Sigma}_u = diag\left(\sigma_1^2, \sigma_2^2, ..., \sigma_N^2\right)$ *and* $\Lambda = (\lambda_1, ..., \lambda_N)'$*, the Kullback-Leibler divergence of* $f(\mathbf{y}_{k,t+1}|s_{t+1} = 1, I_t)$ *with respect* *to* $f(\mathbf{y}_{k,t+1}|s_{t+1}=0,I_t)$ *simplifies to*

$$KL = \frac{(\mu_0-\mu_1)^2}{2}\left(\frac{1}{\sum_{i=1}^{k}\frac{\lambda_i^2}{\sigma_i^2}}+\phi^2 P_{t|t}+\sigma_a^2\right)^{-1}, \tag{14}$$

where $P_{t|t}=E\left(f_t-f_{t|t}|I_t\right)^2$.

Proof: The proof is given in the Appendix.

This expression implies that if there are separate business cycles regimes in the sense that $\mu_1 \neq \mu_0$ and $\sigma_a^2 < \infty$, and the common factor is not weak (as in Lam, Yao and Bathia, 2011) with respect to the early published indicators, that is $\mathbf{\Lambda}_k \neq \mathbf{0}_{k\times 1}$, the divergence is strictly positive. This implies that the incoming information provided by a new release of k economic indicators is always expected to be useful to improve upon the inference about the business cycle at time $t+1$ with respect to using only the information up to t.

In addition, the expression quantifies the magnitude of the change in the conditional entropy and reveals that the averaged gains of using the business cycle information content of $\mathbf{y}_{k,t+1}$ will depend on several parameters. First, the information content of $\mathbf{y}_{k,t+1}$ increases with the difference between the within-state means, $\mu_0-\mu_1$, because the states are more easy to identify from the information provided by the new data. Second, the expression shows that the greater the signal signal-to-noise ratios, $\frac{\lambda_i^2}{\sigma_i^2}$, the larger the expected gains to infer the business cycle regimes from observing $\mathbf{y}_{k,t+1}$. Third, the divergence diminishes with the variance of the common factor, σ_a^2.

Finally, the conditional entropy decreases as $|\phi|$ increases because $\phi^2 P_{t|t}$ is always positive. Therefore, the serial dynamics in the common factor make harder the identification of the business cycle phases because this makes the common factor more variable relative to the difference in means across regimes. Also, $P_{t|t}$ is increasing in ϕ (see, Poncela and Ruiz, 2015) and, therefore, the more persistent the common factor is (in the sense that $\phi \to 1$), more difficult it is to separate expansions from recessions. Then, for a better identification of the business cycle phases, from now on we will assume that $p = 0$ in (3).

In the case of idiosyncratic errors with serial correlation, the integral in (12) does not have a closed form solution since it involves computing the Kullback-Leibler divergence of mixtures of Gaussians (see, for instance, Michalowicz, Nichols and Bucholtz, 2008, or

Cui, 2016, among many others). We will illustrate the value of new information through several simulations and an empirical application in the following sections.

3 Extending the single-index MS-DFM

3.1 State space representation

In what follows, we will focus in the case of interest in our empirical application, the model with just one common factor representing the aggregate economic activity. If we assume that all indicators were always observed with no reporting lags and that the quarterly indicators were sampled each month, the model can be cast into a state space representation

$$\mathbf{y}_t = \mathbf{H}\mathbf{h}_t + \boldsymbol{\eta}_t, \tag{15}$$

$$\mathbf{h}_t = \mathbf{M}_{s_t} + \mathbf{F}\mathbf{h}_{t-1} + \boldsymbol{\xi}_t, \tag{16}$$

with $\boldsymbol{\eta}_t \sim N(0, \mathbf{R})$ and $\boldsymbol{\xi}_t \sim N(0, \mathbf{Q})$.

Without loss of generalization, consider that only the first indicator is sampled quarterly. If it refers to stocks, it can easily be converted into monthly observations since they simply refer to quantities which are measured at a particular time. If it refers to flow variables, Mariano and Murasawa (2003) propose a time aggregation which is based on the notion that its quarterly growth rates, $y_{1,t}$, are weighted averages of its (assumed to be known) monthly past growth rates, x_t:

$$y_{1,t} = \frac{1}{3}x_t + \frac{2}{3}x_{t-1} + x_{t-2} + \frac{2}{3}x_{t-3} + \frac{1}{3}x_{t-4}. \tag{17}$$

For example, if the idiosyncratic components are autoregressive processes of order $p_i = 1$, and if we assume $p = 0$ in (3), the matrices in the measurement equation (15) are $\mathbf{y}_t = (y_{1,t}, y_{2,t}, \cdots, y_{N,t})'$, $\boldsymbol{\eta}_t = \mathbf{0}_{N\times 1}$, $\mathbf{R} = \mathbf{0}_{N\times N}$, $\mathbf{h}_t = (f_t, ..., f_{t-5}, u_{1,t}, ..., u_{1,t-5}, u_{2,t}, \cdots, u_{N,t})'$, and

$$\mathbf{H} = \begin{pmatrix} \frac{\lambda_1}{3} & \frac{2\lambda_1}{3} & \lambda_1 & \frac{2\lambda_1}{3} & \frac{\lambda_1}{3} & \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & \cdots & 0 & 0 \\ \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_N & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}. \tag{18}$$

In this example, the matrices in the transition equation (16) are $\mathbf{M}_{s_t} = (\alpha_{s_t}, \mathbf{0}'_{(N+9)\times 1})'$,

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_{11} & \mathbf{0}_{5\times 5} & \mathbf{0}_{5\times N-1} \\ \mathbf{0}_{5\times 5} & \mathbf{F}_{22} & \mathbf{0}_{5\times N-1} \\ \mathbf{0}_{N-1\times 5} & \mathbf{0}_{N-1\times 5} & \mathbf{F}_{33} \end{pmatrix}, \tag{19}$$

where

$$\mathbf{F}_{11} = \begin{pmatrix} \mathbf{0}_{1\times 4} & 0 \\ \mathbf{I}_4 & \mathbf{0}_{4\times 1} \end{pmatrix}, \tag{20}$$

$$\mathbf{F}_{22} = \begin{pmatrix} (\psi_1, 0, 0, 0) & 0 \\ \mathbf{I}_4 & \mathbf{0}_{4\times 1} \end{pmatrix}, \tag{21}$$

$\mathbf{F}_{33} = diag\,(\psi_2, ..., \psi_N)$, and $\mathbf{Q} = diag\left(\sigma_a^2, ..., 0, \sigma_1^2, 0, ..., 0, \sigma_2^2, ..., \sigma_N^2\right)$.

The model can be estimated by maximum likelihood as follows. Let $\mathbf{h}_{t|t-1}^{(i,j,k,l,m)}$ be the one-period-ahead forecast of $\mathbf{h}_t$ based on information up to period $t-1$, given the path $(s_{t-4} = i, s_{t-3} = j, s_{t-2} = k, s_{t-1} = l, s_t = m)$ and let $\mathbf{P}_{t|t-1}^{(i,j,k,l,m)}$ be its covariance matrix. The prediction equations become

$$\mathbf{h}_{t|t-1}^{(i,j,k,l,m)} = \mathbf{M}_{s_t} + \mathbf{F}\mathbf{h}_{t-1|t-1}^{(i,j,k,l)}, \tag{22}$$

$$\mathbf{P}_{t|t-1}^{(i,j,k,l,m)} = \mathbf{F}\mathbf{P}_{t-1|t-1}^{(i,j,k,l)}\mathbf{F}' + \mathbf{Q}, \tag{23}$$

where $\mathbf{h}_{t-1|t-1}^{(i,j,k,l)}$ is the estimation of $\mathbf{h}_{t-1}$ at time $t-1$ with information up to time $t-1$ if $(s_{t-4} = i, s_{t-3} = j, s_{t-2} = k, s_{t-1} = l)$ and $\mathbf{P}_{t-1|t-1}^{(i,j,k,l)}$ its covariance matrix, that will be given in (27) and (28), respectively. The conditional one-step-ahead forecast errors are $\mathbf{u}_{t|t-1}^{(i,j,k,l,m)} = \mathbf{y}_t - \mathbf{H}\mathbf{h}_{t|t-1}^{(i,j,k,l)}$ and $\mathbf{\Sigma}_{t|t-1}^{(i,j,k,l,m)} = \mathbf{H}\mathbf{P}_{t|t-1}^{(i,j,k,l)}\mathbf{H}' + \mathbf{R}$ is its conditional variance. Hence, the log likelihood given $(s_{t-4} = i, s_{t-3} = j, s_{t-2} = k, s_{t-1} = l, s_t = m)$ can be computed at each t as

$$l_t^{(i,j,k,l,m)} = -\frac{1}{2}\ln\left(2\pi\left|\mathbf{\Sigma}_{t|t-1}^{(i,j,k,l,m)}\right|\right) - \frac{1}{2}\mathbf{u}_{t|t-1}^{(i,j,k,l,m)\prime}\left(\mathbf{\Sigma}_{t|t-1}^{(i,j,k,l,m)}\right)^{-1}\mathbf{u}_{t|t-1}^{(i,j,k,l,m)}. \tag{24}$$

The updating equations become

$$\mathbf{h}_{t|t}^{(i,j,k,l,m)} = \mathbf{h}_{t|t-1}^{(i,j,k,l,m)} + \mathbf{K}_t^{(i,j,k,l,m)}\mathbf{u}_{t|t-1}^{(i,j,k,l,m)}, \tag{25}$$

$$\mathbf{P}_{t|t}^{(i,j,k,l,m)} = \mathbf{P}_{t|t-1}^{(i,j,k,l,m)} - \mathbf{K}_t^{(i,j,k,l,m)}\mathbf{H}\mathbf{P}_{t|t-1}^{(i,j,k,l,m)}, \tag{26}$$

where the Kalman gain, $\mathbf{K}_t^{(i,j,k,l,m)}$, is defined as $\mathbf{K}_t^{(i,j,k,l,m)} = \mathbf{P}_{t|t-1}^{(i,j,k,l,m)}\mathbf{H}'\left(\boldsymbol{\Sigma}_{t|t-1}^{(i,j,k,l,m)}\right)^{-1}$.

In addition, maximizing the exact log likelihood function of the associated nonlinear Kalman filter is computational burdensome since at each iteration, the filter produces a 2-fold increase in the number of cases to consider. The solution adopted in this paper is based on collapsing some terms of the former filter as proposed by Kim (1994) and used by Kim and Yoo (1995) and Chauvet (1998):

$$\mathbf{h}_{t|t}^{(j,k,l,m)} = \frac{\sum_{i=0}^{1} p\left(s_{t-4}=i, s_{t-3}=j, s_{t-2}=k, s_{t-1}=l, s_t=m|I_t\right)\mathbf{h}_{t|t}^{(i,j,k,l,m)}}{p\left(s_{t-3}=j, s_{t-2}=k, s_{t-1}=l, s_t=m|I_t\right)}, \quad (27)$$

$$\mathbf{P}_{t|t}^{(j,k,l,m)} = \frac{\sum_{i=0}^{1} p\left(s_{t-4}=i, s_{t-3}=j, s_{t-2}=k, s_{t-1}=l, s_t=m|I_t\right)\boldsymbol{\Delta}_{t|t}^{(i,j,k,l,m)}}{p\left(s_{t-3}=j, s_{t-2}=k, s_{t-1}=l, s_t=m|I_t\right)}, \quad (28)$$

where $\boldsymbol{\Delta}_{t|t}^{(i,j,k,l,m)} = \mathbf{P}_{t|t}^{(i,j,k,l,m)} + \left(\mathbf{h}_{t|t}^{(j,k,l,m)} - \mathbf{h}_{t|t}^{(i,j,k,l,m)}\right)\left(\mathbf{h}_{t|t}^{(j,k,l,m)} - \mathbf{h}_{t|t}^{(i,j,k,l,m)}\right)'$.

3.2 Missing data

In practice, imposing that all variables are always observed is quite unrealistic. Performing real-time inferences on the business cycle typically requires exploitation of the information of early available indicators, which could be sampled at different frequencies, mainly quarterly and monthly. The presence of ragged-edge data generate missing data at the end of the sample while the missing data appear in quarterly indicators because their figures are only available once every three monthly outcomes.

In linear contexts, Mariano and Murasawa (2003) show that the system of state-space equations remains valid with missing data after a subtle transformation. These authors, fill in the missing observations with random numbers that are extracted from a random variable whose distribution is independent of the model parameters. Then, the measurement equation is modified to get that the Kalman filter skips the random numbers. From a computational point of view, the parameters that maximize the likelihood and the inferences about the business cycle states are achieved as if all the variables were observed.

The procedure can be adapted to Markov-switching models easily. Without loss of generalization, let us focus on dealing with missing data only in the t-th observation of one monthly indicator, $y_{N,t}$, which is the last component of the vector of time series $\mathbf{y}_t$. Let α be the vector that includes all the unknown model parameters. Let us define the variable

$$y_{i,t}^{+} := \begin{cases} y_{i,t} & \text{for } i = 1, ..., N-1 \\ z_t & \text{for } i = N \end{cases}, \tag{29}$$

where z_t is a random variable whose distribution is independent of α and s_t; for instance, $z_t \sim N(0, \sigma_z^2)$. Let $f(z_t; \beta)$ be the density function of z_t, which depends on a vector of parameters β.

Accordingly, no modifications of the nonlinear algorithm used to estimate MS-DFM are required, apart from considering a time varying Kalman filter to zero out the missing observations. Let R_{ii} be the variance of the i-th element of $\mathbf{y}_t$, let $\mathbf{H}_i$ be the i-th row of the matrix $\mathbf{H}$ which has ς columns, and let $\mathbf{0}_{1\varsigma}$ be a row vector of ς zeroes. In our example, the measurement equation can be replaced by the following expressions

$$\mathbf{H}_{it}^{+} := \begin{cases} \mathbf{H}_i & \text{for } i = 1, ..., N-1 \\ 0_{1\varsigma} & \text{for } i = N \end{cases}, \tag{30}$$

$$\boldsymbol{\eta}_{it}^{+} := \begin{cases} 0 & \text{for } i = 1, ..., N-1 \\ z_t & \text{for } i = N \end{cases}, \tag{31}$$

$$\mathbf{R}_{iit}^{+} := \begin{cases} 0 & \text{for } i = 1, ..., N-1 \\ \sigma_z^2 & \text{for } i = N \end{cases}. \tag{32}$$

This trick leads to a time-varying state space model with no missing observations so the nonlinear filter can be directly applied to $\mathbf{y}^{+}$, $\mathbf{H}^{+}$, $\boldsymbol{\eta}_t^{+}$, and $\mathbf{R}^{+}$.

4 Monte Carlo simulations

We generate a total of $M = 1000$ sets of N idiosyncratic components $\mathbf{u}_t^m$ of length T, where $T = 600$, which is about the lag length of the monthly indicators used in our empirical application. Without loss of generality, the time series are generated with equal variances $\sigma_i^2 = \sigma^2$ across time series. However, to examine the effect of the quality of the indicators in the forecasting accuracy, the series are generated with the same but increasing idiosyncratic variance σ^2 of 0.5, 1.5, and 4.5. The dynamics of these idiosyncratic components follow autoregressive processes of order one with autoregressive parameters equal to $\psi_{i,1} = \psi = 0.3$ for all time series. To evaluate the effect of the persistence on the inferences, we also perform simulations with $\psi = 0.6$.

In addition, we generate $M = 1000$ dummy variables b_t^m of zeroes and ones of length T which are used to simulate different sequences of expansions ($b_t^m = 0$) and recessions ($b_t^m = 1$). To ensure that the dummies share the US business cycle properties, b_t^m follows Markov chains with $p_{00} = 0.98$ and $p_{11} = 0.9$.[6] Then, we generate $M = 1000$ common factors, f_t^m, that follow Markov-switching processes given by (3) with $p = 0$. In this case, the business cycle sequences b_t^m to classify the business cycle states and the within state means are set to $\mu_0 = 1$ and $\mu_1 = -1$. Then, to examine how the difference between the within-state means affects the results, μ_0 is increased to 2. Finally, setting $\sigma_a^2 = 1$ and using factor loadings equal to one for all the series, we add the idiosyncratic components to the switching mean factors to generate $M = 1000$ sets of N time series $\{\mathbf{y}_t^m\}_{t=1}^T$.

To examine the effects of dealing with ragged-edge data in computing the real-time business cycle inferences, we assume that an analyst faces the forecasting problem with one publication lag in four out of the set of N indicators used in the analysis. For completeness, the simulations are also computed when these four indicators exhibit two publication lags and the role of N is addressed by using a total number of indicators of 5 and 7.

We consider that the analyst wants to infer the probability of recession at T from the set of N indicators under two different scenarios. The first scenario consists of using traditional MS-DFM to infer recession probabilities at T with the (as large as possible) amount of information disposable at T. In this case, the forecasts are computed from the latest available balanced panel of N indicators. Hence, she has to compute one-step-ahead forecasts to obtain $prob(s_T = 1|I_{T-1})$ and two-step-ahead forecasts to obtain $prob(s_T = 1|I_{T-2})$ from the set of N indicators when there are one and two periods of publication lags, respectively.

The second scenario consists of using our extension of MS-DFM that is able to deal

[6] According to the NBER Business Cycle Dating Committee, these transition probabilities coincide with the percentage of months classified as expansions that are followed by expansions and the percentage of quarters classified as recessions that are followed by recessions in the period used in the empirical application 1967.02-2017.03.

with ragged-edge data. In this case, the inferences can be computed from the set of N indicators even when four of them are not available at T, i.e., the analyst can compute $prob(s_T = 1|I_T^+)$, where I_T^+ refers to the information provided by the set of $N-4$ promptly published indicators up to T and the 4 delayed indicators up to $T-h$, with $h = 1, 2$, when the publication lag is of one and two months, respectively. In this case, the variance of the $N-4$ indicators that are published timely is 1.5, and the variance of the 4 delayed indicators is allowed to change from 0.5 to 1.5, and 4.5

For each m-th replica, we quantify the ability of these procedures to detect the actual state of the business by computing the Forecasting Quadratic Probability Score ($FQPS$):

$$FQPS_i = \frac{1}{M}\sum_{m=1}^{M}(p_{T,i}^m - b_T^m)^2. \tag{33}$$

In this expression, $i = I$ in the case of traditional MS-DFM that forecast from the latest available balanced panel, and $p_{T,i}^m = prob(s_T = 1|I_{T-1})$ or $p_{T,i}^m = prob(s_T = 1|I_{T-2})$ in the cases of one-step-ahead or two-step-ahead forecasts. By contrast, $i = II$ in the case of our extension of MS-DFM, which is able to deal with missing observations, and $p_{T,i}^m = prob(s_T = 1|I_T^+)$. Hence, the measure is the average over the M replications of the squared deviation of the different types of inferences from the generated business cycles.

Table 1 displays the $FQPS$ statistics when four indicators exhibit one and two (in brackets) publication lags. Notably, the table shows that using the incoming information as it is available always helps to increase the accuracy of the models. For example, let us focus on the case of computing inferences from $N = 5$ indicators when four of them exhibit a one-period publication delay in the case $\sigma^2 = 1.5$. When the one-step ahead probability forecasts are computed from the balanced set of five indicators with one lag of publication delay, the inferences computed from the traditional MS-DFM exhibit $FQPS_I$ of 0.069. However, the the MS-DFM that allows ragged-edge data uses one timely available indicator and four indicators with one publication lag to substantially improve the business cycle inferences, reaching a fall in the $FQPS_{II}$ to 0.055. In addition, the accuracy gains of accounting for ragged-edge data increase when the publication delay is two months ($FQPS_I$ of 0.089 vs $FQPS_{II}$ of 0.062, in brackets).

Notably, the sharp increases in the forecasting accuracy detected below are achieved by using only one timely published indicator. When the number of promptly available indicators increases, the inferences computed from the model that accounts for ragged-edge data also outperform those computed from the model that computes probability forecasts from the balanced sets of indicators ($FQPS_I$ of 0.064 vs $FQPS_{II}$ of 0.053), especially when the indicators exhibit larger publication delays ($FQPS_I$ of 0.088 vs $FQPS_{II}$ of 0.056).

The entries displayed in Table 1 show that the ability to compute business cycle inferences from unbalanced panels crucially depends on the signal-to-noise ratio of the early available indicators. Regardless of the forecasting scenario, $FQPS$ rises when the variance of the idiosyncratic component increases from 0.5 to 1.5 and 4.5, and the relative gains of using unbalanced panels diminish as the signal-to-noise ratio increases. In addition, the table shows that higher differences of within-state means, from $\mu_0 = 1$ to $\mu_0 = 2$ for a fixed $\mu_1 = -1$, improve the performance of the models to compute business cycle inferences. Interestingly, the improvements are significantly higher for the model that uses timely available indicators. Finally, the table also shows that the persistence of the idiosyncratic component of the time series tends to diminish the performance of the models.

5 Empirical application

The purpose of this section is to examine the relative empirical performance of our modified MS-DFM, which is able to deal with ragged-edge data and mixed frequencies, with respect to traditional MS-DFM, which are restricted to use balanced panels of data.

5.1 In-sample analysis

The four monthly indicators used in the empirical analysis are industrial production index, nonfarm payroll employment, personal income less transfer payments and real manufacturing and trade sales. Although the latest available data set was downloaded on June, 15th 2017, the balanced panel of four monthly indicators only includes data from 1967.01 to 2017.03, because income is only available up to April 2017 and sales is only available up to March 2017.

Since the seminal proposal of Diebold and Rudebusch (1996), the behavior of these series is assumed to follow the comovements and asymmetries that Burns and Mitchell (1946) designated as the key business cycle features. Following their lines, we fit a MS-DFM to the balanced panel of one hundred times the change in the natural logarithm of these four macroeconomic variables.[7] The maximum likelihood estimates of this monthly model, which are displayed in the top panel of Table 2, show that the estimates of the signal-to-noise ratios agree with the magnitudes used in the simulation experiments. In particular, the highest values of the signal-to-noise ratios are achieved by industrial production, the medium values by employment, and the lowest values by sales and income. In addition, the estimates show that the factor loadings are positive and statistically significant. Hence, the indicators are positively correlated with the estimated common factor.

In line with this statement, Figure 1 shows that the coincident index describes a behavior that closely agrees with the NBER-designated US business cycles.[8] In terms of ROC classification, the common factor capture the state of the US business cycle with AUROC values fairly close to the near-perfect classification ability value of one, with AUROC of 0.94. The magnitude of the AUROC is comparable to that obtained by Berje and Jorda (2011) for two well-known indices of the US business conditions: the Chicago Fed National Activity Index (CFNAI) and the Aruoba, Diebold, and Scotti (2009) Business Conditions Index maintained by the Federal Reserve Bank of Philadelphia.[9]

Notably, the maximum likelihood estimates reported in the top panel of Table 2 also show that the transition probabilities are very persistent ($p_{00} = 0.98$, $p_{11} = 0.85$) and that the within-state means are separated from each other ($\mu_0 = 0.32$, $\mu_1 = -2.00$). Panel A of Figure 1, which plots the probabilities that the coincident indicator is in recession based on currently available information along with shaded areas that represent periods

[7] In line with Stock and Watson (1991), all the linear autoregressive processes are estimated with two lags. Folloing Camacho and Perez Quiros (2007), the nonlinear factor is estimated with no lags.

[8] In the empirical analysis, we take it as given that the NBER correctly identifies the dates of business cycle turning points.

[9] The common factor of a linear DFM has AUROC=0.94. Although MS-DFM exhibits similar performance in terms of ROC classification, it has the advantage of computing the probability that the common factor is in recession.

dated as recessions by the NBER, shows that the smoothed probabilities are in striking agreement with the professional consensus as to the history of US business cycles.

Our extension of MS-DFM allows us to obtain nonlinear estimates and business cycle inferences from a data set that contains business cycle indicators with monthly and quarterly frequencies. In addition to the four monthly indices, we include the growth rate of U.S. quarterly real GDP from 1967.1 to 2017.1. The bottom panel of Table 2 shows that the maximum likelihood estimates that refer to the monthly series and the common factor are similar to the estimates obtained when GDP was not included in the model. The dynamics of the common factor, which is plotted in Panel B of Figure 1, is also in close agreement with the dynamics of the estimated common factor obtained from the model that excludes GDP. An interesting result is that its AUROC is also 0.94, which reveals that quarterly GDP does not seem to improve the in-sample classification ability of the four monthly indicators.

5.2 Real-time analysis

The previous in-sample analysis has been conducted with data of the most recent vintage. However, the real-time data could be less helpful in monitoring the real activity than the in-sample evaluations developed in the previous section using finally revised data sets. On the one hand, it has been argued in the related literature (see, for example, Diebold and Rudebusch, 1991) that the good performance of the end-of-sample vintages in examining the empirical performance of econometric models may be spurious, in the sense that the data actually available in real time include economic time series that are subject to revision and that the economic relationships may change over time. In our case, the measures of production, employment and sales are typically subject to substantial revisions that sometimes occur years after the official figures are firstly released. On the other hand, the in-sample analysis does not allow the researchers to evaluate the effects of managing the lack of synchronicity that characterizes the daily flow of macroeconomic information in the early assessments of the economic developments.

To perform a more realistic assessment of the actual empirical reliability of the MS-DFM, we evaluate their real-time performance at tracking the US business cycles in real time through a data set that consists of real-time vintages obtained from January 15, 1976 to June 15, 2017. That is, the inferences are computed at each month t over the past four decades that covers the period December, 1976 to May, 2011 by using only the data that would have been available at the middle of the month that follows the particular month in which the inference is computed.[10] Hence, the real-time analysis does not include the data revisions that were not available at the time the model would have been used and has to manage with incomplete data sets at the time of each inference.

To clarify understanding, let us describe the stylized publication calendar of the economic indicators used in the real-time analysis. At the end of month t, Industrial Production is published on the 15th of the month $t+1$; Non-farm Employees is published on the 8th of the month $t+1$, Real Personal Income is published on the 27th of the month $t+1$, and Real Manufacturing, Trade Sales is published on the 27th of the month $t+2$.[11] In addition, GDP is published on the 15th of $t+2$, whenever t is March, June, September or December. To simplify the real-time analysis, we consider that the real-time inferences are computed on the 15th of each month, where employment and industrial production are available for the previous month. On this day, of month $t+1$, we infer the probabilities of being in a recession at t, with industrial production and employment up to t, personal income up to $t-1$, and real sales up to $t-2$.

According to our theoretical and Monte Carlo results, the business cycle probabilities are inferred from different alternative strategies. The first two strategies consist of computing inferences from traditional MS-DFM which can only account for balanced data sets. This implies that the model cannot use either quarterly series or the information provided by the early published indicators since the data set must be constrained to finish at $t-2$. Within this strategy, called strategy A, the inferences computed at $t-2$ are considered as the prevailing business cycle conditions for period t, i.e., the probabilities of being in a recession at t are approximated by $[prob(s_{t-2}=1|I_{t-2})]_t$. In the second strategy, called strategy B, the probabilities at t are computed by projecting the estimated probabilities

[10] We use the real-time dataset archived at the Federal Reserve Bank of Philadelphia, which includes the history of all the indicators that would actually have been available to a researcher at any given point in time.

[11] The nominal indicator is published on the 14th of $t+2$.

for period $t-2$ to the current state by multiplying latest inferences by the transition matrix, $prob(s_t = 1|I_{t-2})$.

Strategies A and B clearly miss the extremely valuable information about the current business cycle that is provided by the early published indicators. In particular, these inferences miss the data of personal income at $t-1$, and industrial production and employment at $t-1$ and t. To overcome this drawback, the business cycle inferences are computed in strategy C by using the extension of MS-DFM proposed in this paper to deal with ragged-edge data. Finally, the inferences are also computed in strategy D by enlarging this model with GDP, which requires dealing with mixed frequencies described in Section 2.

Figure 2, which plots the real-time filtered probabilities estimated from strategies A (panel A), B (panel B), C (panel C), and D (panel D), respectively, helps us to assess the empirical performance of the different strategies in real time. As expected, when the analysis is developed in real time the figures show a significant deterioration in the models' performance with respect to the in-sample results. Although the in-sample filtered probabilities plotted in Figure 1, which are computed from finally revised data, provide unequivocal jumps in probabilities that marked the start and the end of the US business cycle phases, the real-time probabilities plotted in Figure 2 produce noisier and less accurate signals of the business cycle.

The figures also show that there is a significant improvement in business cycle forecasting accuracy when the MS-DFM is allowed to deal with ragged-edge data. To evaluate these forecasting improvements, Table 3 displays the results of the real-time $FQPS$ for the four different strategies. According to our theoretical and Monte Carlo results, strategy C provides much better forecasting accuracy than strategies A and B, with a reduction in $FQPS$ of more than 35%. To analyze whether empirical loss differences across the competing models and Strategy C are statistically significant, the last row of the table shows in brackets the p-values of the pairwise test introduced by Diebold and Mariano (DM, 1995), which is the most influential test in the literature of equal forecasting accuracy. According to the p-values of the DM test, the reductions in $FQPS$ achieved with the model that manage ragged-ends data are statistically significant at standard significance levels.

Notably, the MS-DFM that additionally accounts for mixed frequencies does not seem to exhibit significant reductions in $FQPS$ to that achieved by the MS-DFM that uses the unbalanced panel of only four monthly indicators ($FQPS$ of 0.055 and 0.052, respectively). In addition, the equal predictive accuracy tests show that the differences in forecasting accuracy between Strategy C and Strategy D are not statistically significant. Therefore, our empirical application shows that, although using early available information from the set of four monthly indicators substantially improves the US business cycle inferences in real time, adding the information provided by quarterly GDP does not appear to be relevant.

6 Concluding remarks

Real-time data usually display the feature of ragged ends, which means that end-of-sample observations of time series are missing and only released with a time-lag. The asynchronous publication releases limit the empirical benefits of Markov-switching dynamic factor models in monitoring the day-to-day economic developments because these models are restricted to dealing with balanced data vintages and cannot manage all the relevant new releases as they arrive. In practice, the business cycle inferences computed from these models are either available only with a delay of several months or they are computed as forecasts of past inferences.

From the point of view of monitoring business cycle conditions, we show in the paper that there is no reason to be late or to disregard the relevant information provided by the latest figures of promptly issued indicators. We theoretically show that, when the economic indicators are carefully selected to have large signal-to-noise ratios in the Kalman filter used to compute business cycle inferences, the increase in the accuracy of business cycle identification becomes substantial.

The extension of dynamic factor models with regime switches proposed in this paper is the missing piece of this puzzle. Following the linear proposal of Mariano and Murasawa (2003), the method is based on a nonlinear Kalman filter to fill in the gaps of the non-synchronous flow of data releases in an efficient manner. By means of several Monte Carlo experiments, we quantify the magnitude of the accuracy improvements provided by our proposal over traditional methods, which substantially depends on the signal-to-noise ratio of the early available indicators.

In addition, traditional Markov-switching dynamic factor models cannot deal with business cycle indicators of different -typically monthly and quarterly- frequencies. In this paper, we also show how to mix monthly and quarterly indicators to infer the business cycle phases. The method treat quarterly data as monthly data that exhibit missing monthly observations within each quarter. Accordingly, the nonlinear state-space framework proposed to deal with ragged-edge data can also be used to combine business cycle indicators of different frequencies.

In the empirical application considered in this paper, we find that our theoretical findings are borne out. We use a real-time collection of data vintages which are updated monthly using only the information that would have been available at each month over the last four decades. The vintages use the four constituent monthly series of the Stock-Watson coincident index. Our extension produces real-time business cycle probabilities that track the business cycle accurately, with pronounced drops corresponding to the NBER-designated recessions. Notably, we obtain substantial improvements in our extension of Markov-switching dynamic factor models with respect to forecasting the probabilities from balanced panels of indicators. However, we failed to find significant improvements when GDP is used as an additional indicator.

One potential limitation of the model is that it does not take into account revisions to the indicator initial releases, which are known to be sometimes quite large. The inaccuracy of initial data complicates decision making by policymakers and other agents whose optimal choices depend on the state of the economy. We consider that this extension is important enough to leave it for further research.

Appendix

Proof of Proposition 1: To evaluate the information content of $\mathbf{y}_{k,t+1|t}$, note that

$$f(\mathbf{y}_{k,t+1|t}|s_{t+1}=i,I_t)=\frac{1}{\left(\sqrt{2\pi}\right)^k\left|\boldsymbol{\Sigma}_{k,t+1|t}^{(i)}\right|^{1/2}}\exp(-\frac{1}{2}(\mathbf{y}_{k,t+1}-\boldsymbol{\Lambda}_k\boldsymbol{f}_{t+1|t}^{(i)})'\left(\boldsymbol{\Sigma}_{k,t+1|t}^{(i)}\right)^{-1}(\mathbf{y}_{k,t+1}-\boldsymbol{\Lambda}_k\boldsymbol{f}_{t+1|t}^{(i)})) \tag{A1}$$

for $i=0,1$where $\boldsymbol{\Sigma}_{k,t+1|t}^{(i)}$ is the variance of $\mathbf{y}_{k,t}|s_{t+1}=i,I_t$, which for brevity we will denote $\boldsymbol{\Sigma}_i$ and $\boldsymbol{f}_{t+1|t}^{(i)}=E(\boldsymbol{f}_{t+1}|s_{t+1}=i,I_t)$ denoted more briefly as $\boldsymbol{f}_i$. Using the notation in Cover and Thomas (2006) for the multiple integral, then

$$\begin{aligned}
KL &= \int \ln\frac{f(\mathbf{y}_{k,t+1}|s_{t+1}=1,I_t)}{f(\mathbf{y}_{k,t+1}|s_{t+1}=0,I_t)}f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1} \\
&= \int -\ln(f(\mathbf{y}_{k,t+1}|s_{t+1}=0,I_t))f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1} \\
&- \int -\ln(f(\mathbf{y}_{k,t+1}|s_{t+1}=1,I_t))f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1} \\
&= K_0-K_1,
\end{aligned}$$

where K_i is the entropy of $f(\mathbf{y}_{k,t+1}|s_{t+1}=i,I_t)$, $i=0,1$ with respect to the conditional density $f(\mathbf{y}_{k,t+1}|s_t=1,I_t)$. The entropy of a Gaussian distribution is given by (see, for instance, Cover and Thomas, 2006)

$$K_1=\frac{1}{2}\ln(2\pi)^k\left|\boldsymbol{\Sigma}_1\right|+\frac{k}{2}.$$

The first part of the integral K_0 is given by

$$\begin{aligned}
K_0 &= \int -\ln(f(\mathbf{y}_{k,t+1}|s_{t+1}=0,I_t))f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1} \\
&= \frac{1}{2}\int(\mathbf{y}_{k,t+1}-\boldsymbol{\Lambda}_k\boldsymbol{f}_0)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{y}_{k,t+1}-\boldsymbol{\Lambda}_k\boldsymbol{f}_0)f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1} \\
&+ \frac{1}{2}\int\ln(2\pi)^k\left|\boldsymbol{\Sigma}_0\right|f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1}
\end{aligned}$$

Summing and subtracting $\boldsymbol{\Lambda}_k\boldsymbol{f}_1$ so the expectation in the previous expression does not change, the previous integral can be written as

$$\begin{aligned}
K_0 &= \int -\ln(f(\mathbf{y}_{k,t+1}|s_{t+1}=0,I_t))f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1} \\
&= \frac{1}{2}\int(\mathbf{y}_{k,t+1}-\boldsymbol{\Lambda}_k\boldsymbol{f}_1-\boldsymbol{\Lambda}_k\boldsymbol{f}_0+\boldsymbol{\Lambda}_k\boldsymbol{f}_1)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{y}_{k,t+1}-\boldsymbol{\Lambda}_k\boldsymbol{f}_1-\boldsymbol{\Lambda}_k\boldsymbol{f}_0+\boldsymbol{\Lambda}_k\boldsymbol{f}_1)f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1} \\
&+ \frac{1}{2}\ln(2\pi)^k\left|\boldsymbol{\Sigma}_0\right| \\
&= \frac{1}{2}\int(\mathbf{y}_{k,t+1}-\boldsymbol{\Lambda}_k\boldsymbol{f}_1)'\boldsymbol{\Sigma}_0^{-1}(\mathbf{y}_{k,t+1}-\boldsymbol{\Lambda}_k\boldsymbol{f}_1)f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1} \\
&+ \frac{1}{2}\int(\boldsymbol{f}_1-\boldsymbol{f}_0)'\boldsymbol{\Lambda}_k'\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Lambda}_k(\boldsymbol{f}_1-\boldsymbol{f}_0)f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1} \\
&- \int(\mathbf{y}_{k,t+1}-\boldsymbol{\Lambda}_k\boldsymbol{f}_1)'\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Lambda}_k(\boldsymbol{f}_1-\boldsymbol{f}_0)f(\mathbf{y}_{k,t+1}|s_t=1,I_t)d\mathbf{y}_{k,t+1} \\
&+ \frac{1}{2}\ln(2\pi)^k\left|\boldsymbol{\Sigma}_0\right| \\
&= \frac{1}{2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1}+(\boldsymbol{f}_1-\boldsymbol{f}_0)'\boldsymbol{\Lambda}_k'\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Lambda}_k(\boldsymbol{f}_1-\boldsymbol{f}_0)+\frac{1}{2}\ln(2\pi)^k\left|\boldsymbol{\Sigma}_0\right|
\end{aligned}$$

where to obtain the result we have taken into account that $E(\mathbf{y}_{k,t+1}|s_{t+1}=1,I_t)=\boldsymbol{\Lambda}_k\boldsymbol{f}_1$. Then, The Kullback Leibler divergence is given by

$$\begin{aligned}
KL &= K_0-K_1 \\
&= \frac{1}{2}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1}+(\boldsymbol{f}_1-\boldsymbol{f}_0)'\boldsymbol{\Lambda}_k'\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Lambda}_k(\boldsymbol{f}_1-\boldsymbol{f}_0)+\frac{1}{2}\ln(2\pi)^k(\left|\boldsymbol{\Sigma}_0\right|-\left|\boldsymbol{\Sigma}_1\right|)-\frac{k}{2},
\end{aligned}$$

which is the desired expression.

Proof of Corollary: In this case

$$f(\mathbf{y}_{k,t+1|t}|s_{t+1}=i,I_t)=\frac{1}{\left(\sqrt{2\pi}\right)^k\left|\mathbf{\Sigma}_{k,t+1|t}\right|^{1/2}}\exp(-\frac{1}{2}(\mathbf{y}_{k,t+1}-\mathbf{\Lambda}_k f^{(i)}_{t+1|t})'\mathbf{\Sigma}^{-1}_{k,t+1|t}(\mathbf{y}_{k,t+1}-\mathbf{\Lambda}_k f^{(i)}_{t+1|t})) \tag{A1}$$

for $i=0,1$where $\mathbf{\Sigma}_{k,t+1|t}$ is the variance of $\mathbf{y}_{k,t}|s_{t+1}=i,I_t$, which we have assumed to be the same in both states and $f^{(i)}_{t+1|t}=E(f_{t+1}|s_{t+1}=i,I_t)$. Then, the Kullback-Leibler divergence is now given by (**??**) for the case $r=1$,

$$\begin{aligned} KL &= \frac{1}{2}\left(f^{(0)}_{t+1|t}-f^{(1)}_{t+1|t}\right)^2\mathbf{\Lambda}'_k\mathbf{\Sigma}^{-1}_{k,t+1|t}\mathbf{\Lambda}_k \\ &= \frac{1}{2}\left(\mu_0-\mu_1\right)^2\mathbf{\Lambda}'_k\mathbf{\Sigma}^{-1}_{k,t+1|t}\mathbf{\Lambda}_k. \end{aligned} \tag{A2}$$

Let $\mathbf{\Sigma}_{k,u}=didag(\sigma_1^2,...,\sigma_k^2)$ be the idiosyncratic variance of the k promptly issued indicators $\mathbf{y}_{k,t+1}$. Taking into account the expression for the inverse of the sum of two matrices (see, for instance, Rao, 1973)

$$\mathbf{\Sigma}^{-1}_{k,t+1|t}=\mathbf{\Sigma}^{-1}_{k,u}-\mathbf{\Sigma}^{-1}_{k,u}\mathbf{\Lambda}_k\left(P^{-1}_{t+1|t}+\mathbf{\Lambda}'_k\mathbf{\Sigma}^{-1}_{k,u}\mathbf{\Lambda}_k\right)^{-1}\mathbf{\Lambda}'_k\mathbf{\Sigma}^{-1}_{k,u}. \tag{A3}$$

where $P_{t+1|t}=E\left(f_{t+1}-f_{t+1|t}|I_t\right)^2$ is given by

$$P_{t+1|t}=\phi^2P_{t|t}+\sigma_a^2. \tag{A4}$$

Plugging (A3) and (A4) into (A2)

$$\begin{aligned} KL &= \frac{1}{2}\left(\mu_0-\mu_1\right)^2\left(\sum_{i=1}^k\frac{\lambda_i^2}{\sigma_i^2}-\frac{\left(\sum_{i=1}^k\frac{\lambda_i^2}{\sigma_i^2}\right)^2}{\sum_{i=1}^k\frac{\lambda_i^2}{\sigma_i^2}+\frac{1}{\phi^2P_{t|t}+\sigma_a^2}}\right) \\ &= \frac{1}{2}\left(\mu_0-\mu_1\right)^2\frac{\frac{1}{\phi^2P_{t|t}+\sigma_a^2}\sum_{i=1}^k\frac{\lambda_i^2}{\sigma_i^2}}{\sum_{i=1}^k\frac{\lambda_i^2}{\sigma_i^2}+\frac{1}{\phi^2P_{t|t}+\sigma_a^2}} \\ &= \frac{1}{2}\left(\mu_0-\mu_1\right)^2\left(\frac{1}{\sum_{i=1}^k\frac{\lambda_i^2}{\sigma_i^2}}+\phi^2P_{t|t}+\sigma_a^2\right)^{-1} \end{aligned} \tag{A5}$$

which is the desired expression.

References

[1] Aruoba, B., Diebold, F., and Scotti, C. 2009. Real-time measurement of business conditions. *Journal of Business and Economic Statistics* 27: 417-427.

[2] Barhoumi, K, Darne O. and Ferrara, L. Dynamic Factor Models: A review of the literature. 2014. Journal of Business Cycle Measurement and Analysis, Number 8, Issue 2

[3] Berje, T, and Jorda, O. 2011. The classification of economic activity into expansions and recessions. *American Economic Journal: Macroeconomics* 3: 246-277

[4] Burns, A., and Mitchell, W. 1946. Measuring business cycles. National Bureau of Economic Research, New York.

[5] Camacho, M., and Perez Quiros, G., and Poncela, P. 2015. Extracting nonlinear signals from several economic indicators. *Journal of Applied Econometrics* 30: 1073-1089.

[6] Camacho, M., and Perez Quiros, G. 2007. Jump-and-rest effect of U.S. business cycles. *Studies in Nonlinear Dynamics and Econometrics* 11(4): article 3.

[7] Chauvet, M. 1998. An econometric characterization of business cycle dynamics with factor structure and regime switches. *International Economic Review* 39: 969-96.

[8] Chauvet, M., and Hamilton, J. 2006. Dating business cycle turning points in real time. In *Nonlinear Time Series Analysis of Business Cycles*, eds. C. Milas, P. Rothman, and D. Van Dijk. Amsterdam: Elsevier Science, pp. 1-54.

[9] 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.

[10] Cover, T. and Thomas, J. 2006. *Elements of information theory.*Wiley, New Jersey.

[11] Cui, S. 2016. Comparison of approximation methods to Kullback-Leibler divergence between Gaussian mixture models for satellite image retrieval. *Remote Sensing Letters* 7(7): 651-660.

[12] Diebold, F., and Rudebusch, G. 1991. Forecasting output with the composite leading index: A real-time analysis. *Journal of the American Statistical Association* 86: 603-610.

[13] Diebold, F., and Rudebusch, G. 1996. Measuring business cycles: A modern perspective. *Review of Economics and Statistics* 78: 67-77.

[14] Evans M. 2005. Where are we now? Real time estimates of the macroeconomy. *International Journal of Central Banking* 1: 127-175.

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

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

[17] Kim, C. 1994. Dynamic linear models with Markov-switching. *Journal of Econometrics* 60: 1-22.

[18] Kim, C., and Nelson, C. 1998. Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. *Review of Economics and Statistics* 80: 188-201.

[19] Kim, C., and Yoo, J.S. 1995. New index of coincident indicators: A multivariate Markov switching factor model approach. *Journal of Monetary Economics* 36: 607-630.

[20] Lam, C., Yao, Q., and Bathia, N. 2011. Estimation of latent factors using high-dimensional time series. *Biometrika*, 98(4):901-918.

[21] Mariano, R., and Murasawa, Y. 2003. A new coincident index os business cycles based on monthly and quarterly series. *Journal of Applied Econometrics* 18: 427-443.

[22] Michalowicz, J.V., Nichols, J.M. and Bucholtz, F. 2008. Calculation of differential entropy for a mixed Gaussian distribution. *Entropy* 10: 200-206.

[23] Poncela, P. and Ruiz, E. 2015. More is not always better: back to the Kalman filter in Dynamic Factor Models in *Unobserved Components and Time Series Econometrics* (Koopman and Shephard, Eds), chapter 11. Oxford University Press.

[24] Proietti, T. and Moauro, F. 2006. Dynamic factor analysis with non-linear temporal aggregation constraints. *Appl. Statist.* 55: 281-300.

[25] Rao, C. R. 1973. *Linear Statistical Inference and Its Applications*. New York: J. Wiley.

[26] Stock, J., and Watson, M. 1991. A probability model of the coincident economic indicators. In *Leading Economic Indicators: New Approaches and Forecasting Records*, edited by K. Lahiri and G. Moore, Cambridge University Press.

Table 1. Analysis of ragged ends in MS-DFM

		Balanced panels				Unbalanced panels			
	σ^2	ψ=0.3		ψ=0.6		ψ=0.3		ψ=0.6	
		N=5	N=7	N=5	N=7	N=1+4	N=3+4	N=1+4	N=3+4
μ_0=1, μ_1=-1	0.5	0.062	0.060	0.069	0.061	0.047	0.043	0.055	0.048
		(0.086)	(0.085)	(0.089)	(0.087)	(0.051)	(0.046)	(0.062)	(0.051)
	1.5	0.069	0.066	0.079	0.069	0.055	0.053	0.063	0.055
		(0.089)	(0.088)	(0.094)	(0.092)	(0.062)	(0.056)	(0.071)	(0.061)
	4.5	0.077	0.076	0.104	0.085	0.070	0.068	0.084	0.072
		(0.099)	(0.098)	(0.109)	(0.106)	(0.085)	(0.079)	(0.088)	(0.073)
	σ^2	ψ=0.3		ψ=0.6		ψ=0.3		ψ=0.6	
		N=5	N=7	N=5	N=7	N=1+4	N=3+4	N=1+4	N=3+4
μ_0=2, μ_1=-1	0.5	0.047	0.041	0.048	0.042	0.028	0.018	0.031	0.020
		(0.073)	(0.072)	(0.074)	(0.073)	(0.031)	(0.022)	(0.035)	(0.028)
	1.5	0.048	0.043	0.052	0.045	0.031	0.021	0.035	0.024
		(0.075)	(0.073)	(0.078)	(0.070)	(0.035)	(0.028)	(0.041)	(0.031)
	4.5	0.064	0.053	0.079	0.062	0.042	0.026	0.054	0.038
		(0.082)	(0.080)	(0.092)	(0.088)	(0.050)	(0.035)	(0.061)	(0.048)

Notes. N is the number of indicators and σ^2 is the variance of their idiosyncratic components. In balanced MS-DFM, entries show the average over the replications of the squared deviation of one- and two- (in brackets) step-ahead filtered probabilities of low-mean state from the 1000 generated business cycle sequences. In MS-DFM with unbalanced panels, 1 and 3 variables with variance 1.5 are assumed to be timely available when the inference is computed and 4 indicators with variance 0.5, 1.5 and 4.5 are published with one- and two-month lags. High-growth and low-growth means are μ_0 and μ_1, respectively; ψ is the autoregressive parameter of the idiosyncratic components.

Table 3. Maximum likelihood estimates

Monthly						
		IP	Empl	Inc	Sales	
Indicators	λ_i	0.69 (0.03)	0.42 (0.02)	0.28 (0.04)	0.46 (0.03)	
	ψ_1	-0.18 (0.08)	0.24 (0.03)	-0.20 (0.02)	-0.34 (0.04)	
	ψ_2	-0.16 (0.08)	0.54 (0.04)	-0.05 (0.04)	-0.15 (0.05)	
	σ_i^2	0.26 (0.04)	0.27 (0.02)	0.85 (0.03)	0.57 (0.03)	
Factor	μ_0	μ_1	$\sigma_{a^*}^2$	p_{00}	p_{11}	
	0.32 (0.07)	-2.00 (0.20)	1	0.98 (0.01)	0.85 (0.05)	
Monthly and quarterly						
		IP	Empl	Inc	Sales	GDP
Indicators	λ_i	0.67 (0.03)	0.42 (0.02)	0.29 (0.04)	0.48 (0.03)	0.30 (0.02)
	ψ_1	-0.07 (0.07)	0.24 (0.03)	-0.21 (0.02)	-0.37 (0.04)	0.024 (0.40)
	ψ_2	-0.07 (0.07)	0.54 (0.04)	-0.06 (0.04)	-0.17 (0.05)	-0.57 (0.22)
	σ_i^2	0.55 (0.04)	0.51 (0.02)	0.91 (0.03)	0.73 (0.03)	0.47 (0.15)
Factor	μ_0	μ_1	$\sigma_{a^*}^2$	p_{00}	p_{11}	
	0.29 (0.07)	-2.00 (0.21)	1	0.98 (0.01)	0.83 (0.06)	

Notes. Standard errors are in parentheses

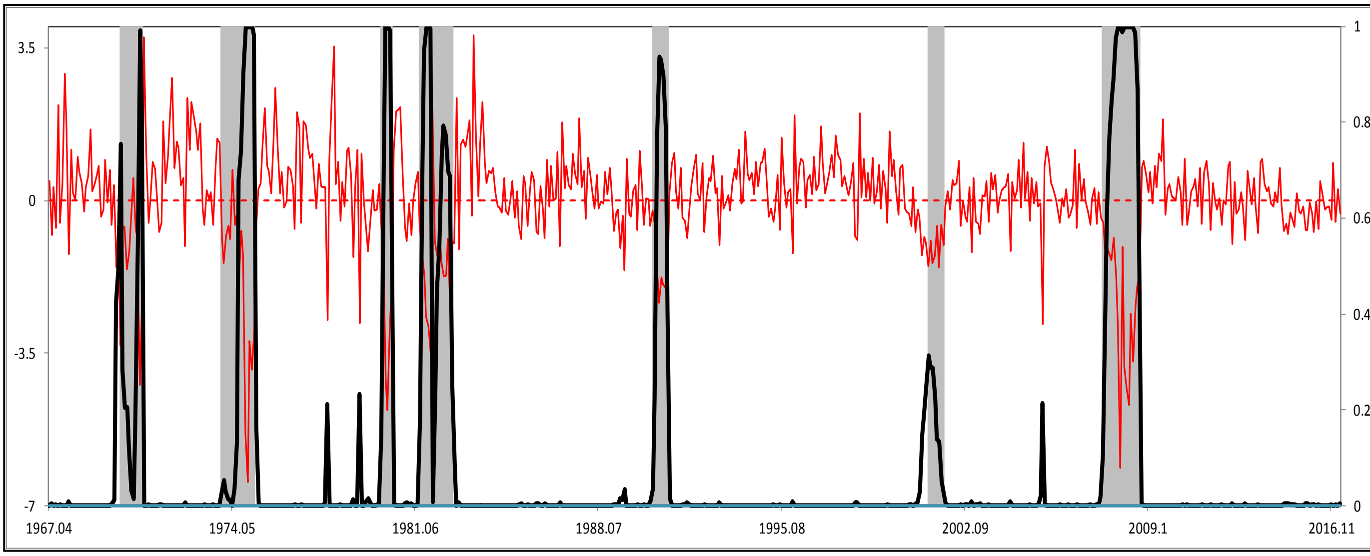
Table 3. Real-time (1976.10-2017.05) empirical performance

Strategy A	Strategy B	Strategy C	Strategy D
Balanced	Balanced	Unbalanced	Unbalanced Mixed frequency
0.086 (0.005)	0.083 (0.006)	0.055 (-)	0.052 (0.321)

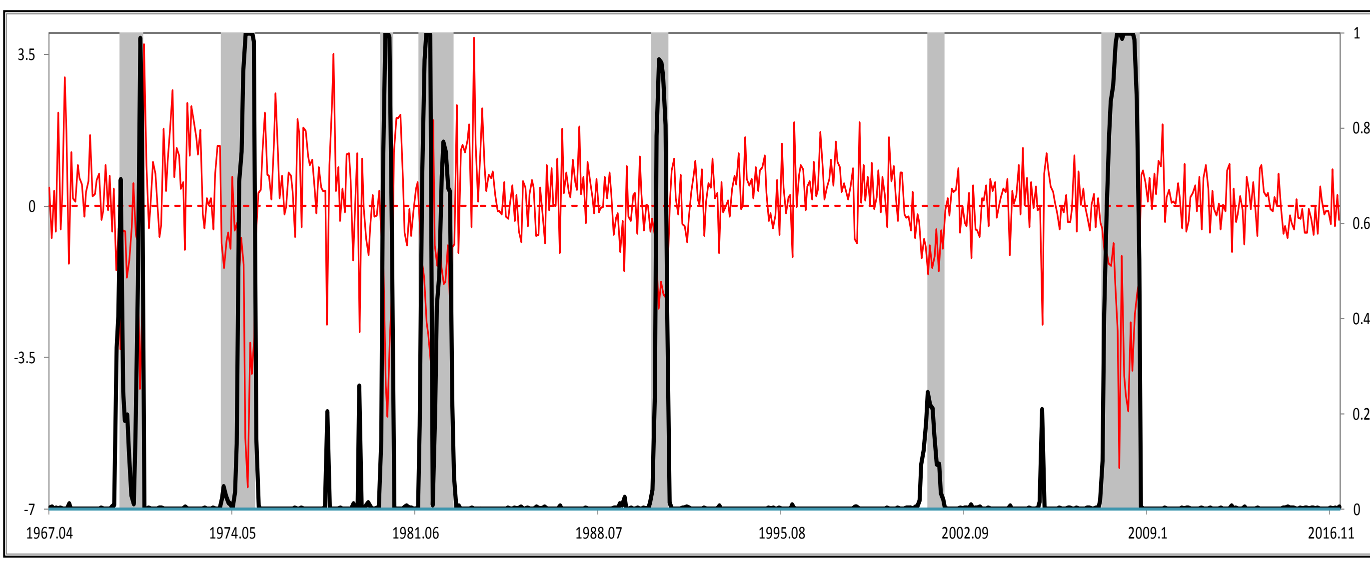
Note. Entries refer to *FQPS* statistics. The figures within parentheses show the *p*-values of the Diebold-Mariano (DM) test of equal forecast accuracy between Strategy C and the rest of strategies. The forecasting strategies are defined in the text.

Figure 1. In-sample results 1967.04-2017.03

Panel A. Four monthly indicators



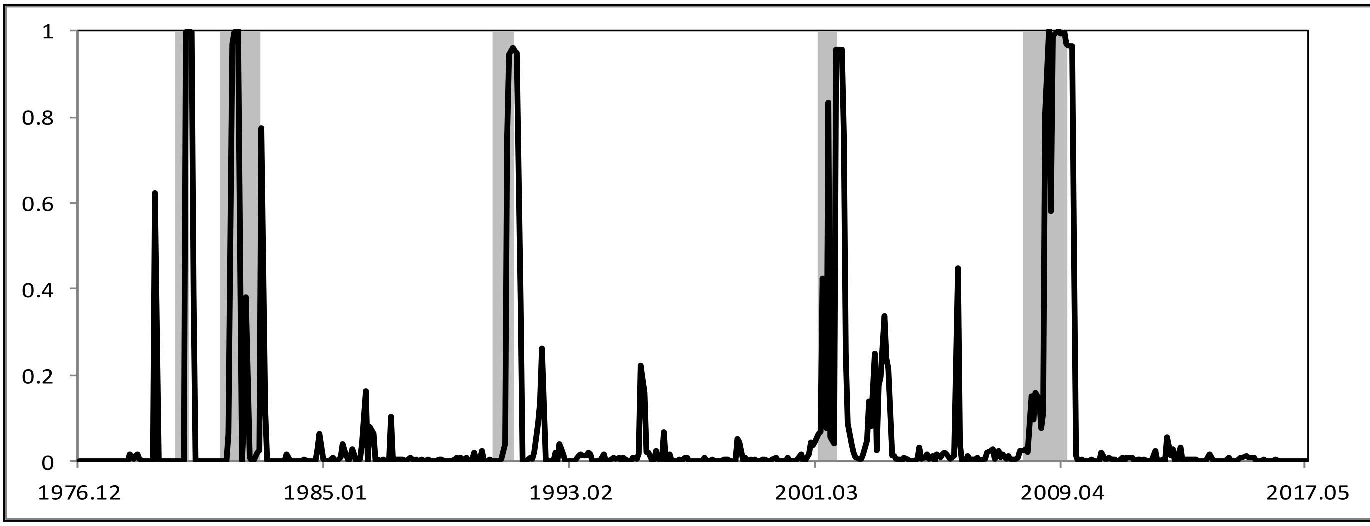
Panel B. Four monthly indicators and quarterly GDP



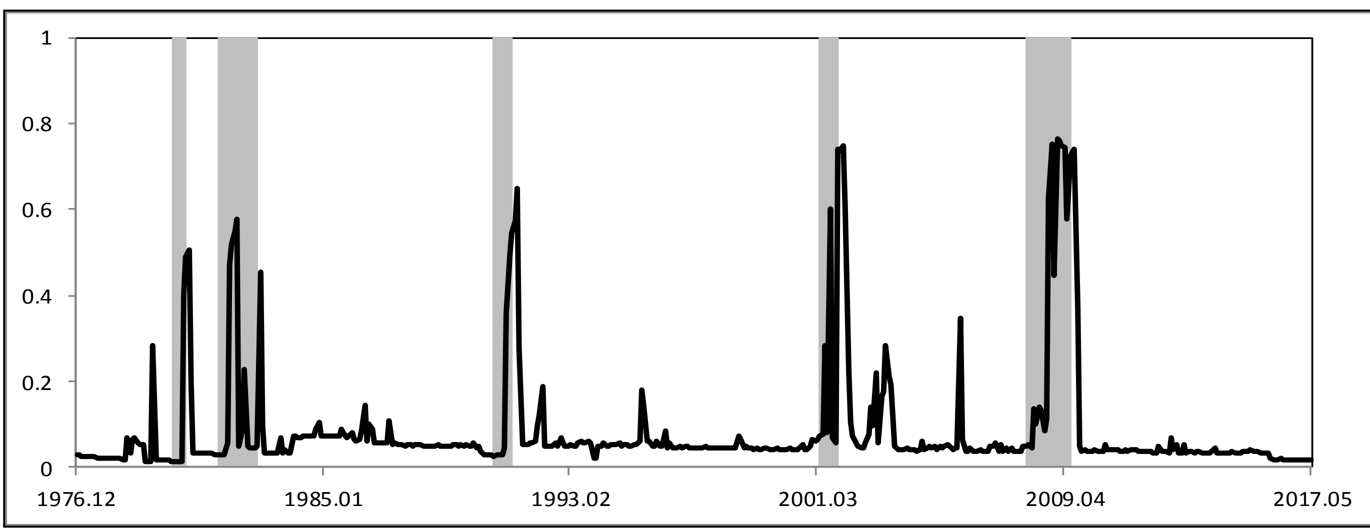
Note. These graph plots the common factor (left scale) and the smoothed probabilities of recession (right scale). Panel A refers to refers to the model of four monthly indicators while Panel B refers to the model that adds GDP. Shaded areas correspond to recessions as documented by the NBER.

Figure 2. Real time probabilities 1976.12-2017.05

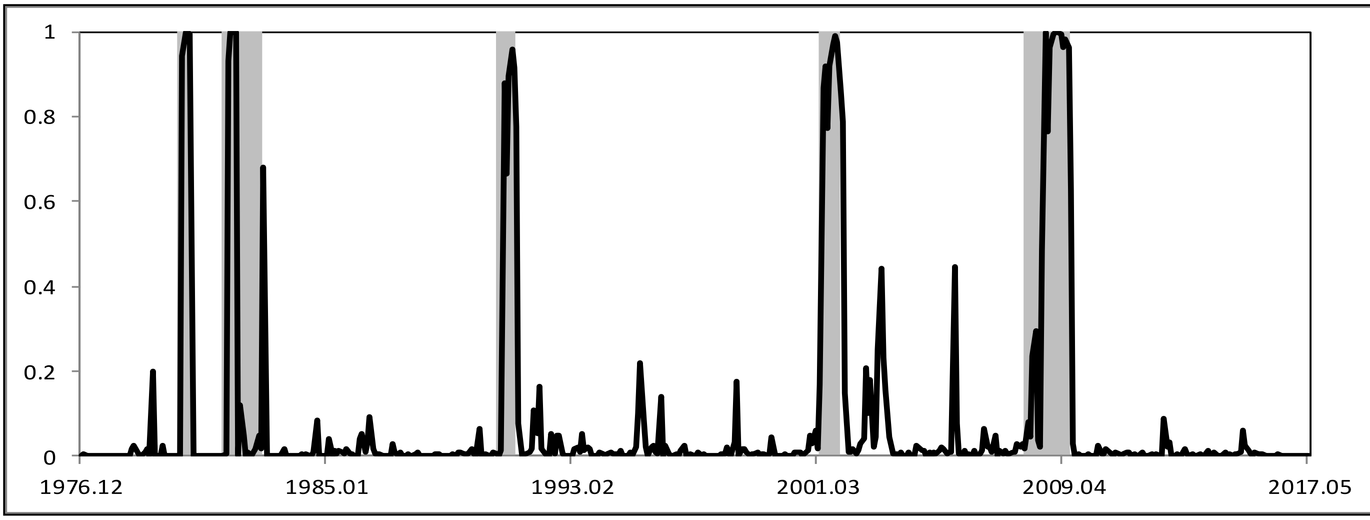
Panel A. Strategy A



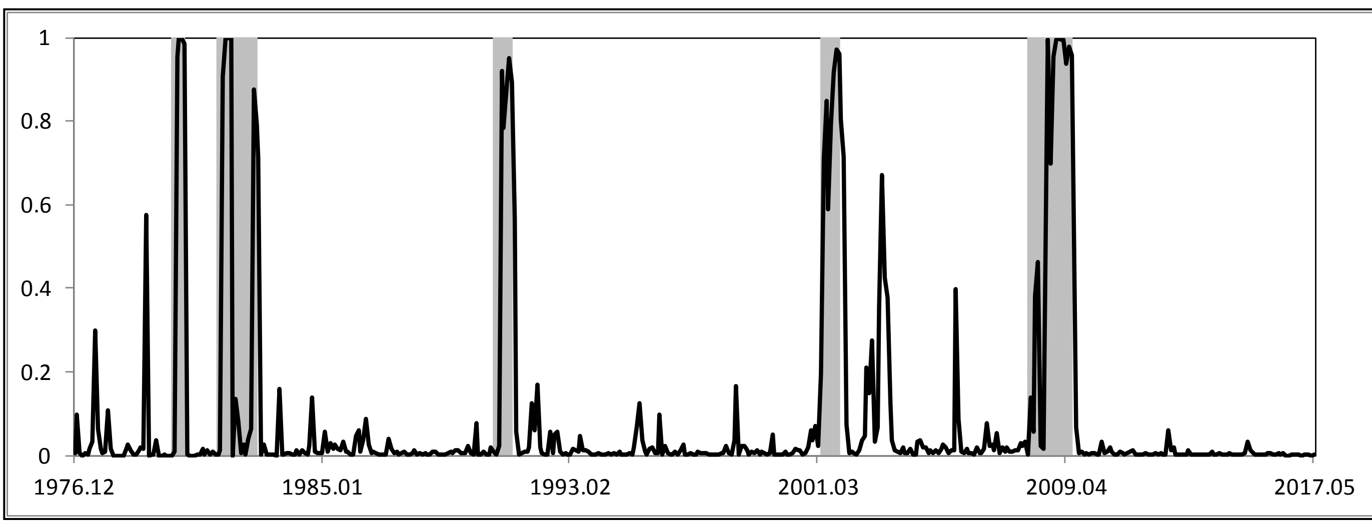
Panel B. Strategy B



Panel C. Strategy C



Panel D. Strategy D



Notes. Panel A reports the probabilities of recession in month *t-2* (estimated using a balanced panel of data) plotted in *t*. Panel B reports the forecasted probabilities of recession in *t* with information up to *t*-2. Panel C reports the probabilities of recession in *t* estimated with an unbalanced panel up to *t*. In addition, Panel D plots the probabilities from a model that also accounts for mixed frequencies. Shaded areas correspond to recessions as documented by the NBER.