Extracting nonlinear signals from several economic indicators.[*]

Maximo Camacho[†]

Universidad de Murcia

mcamacho@um.es

Gabriel Perez-Quiros

Banco de España and CEPR

gabriel.perez@bde.es

Pilar Poncela

Universidad Autónoma de Madrid

pilar.poncela@uam.es

June 17, 2014

Abstract

We develop a twofold analysis of how the information provided by several economic indicators can be used in Markov-switching dynamic factor models to identify the business cycle turning points. First, we compare the performance of a fully non-linear multivariate specification (one-step approach) with the "shortcut" of using a linear factor model to obtain a coincident indicator which is then used to compute the Markov-switching probabilities (two-step approach). Second, we examine the role of increasing the number of indicators. Our results suggest that one step is generally preferred to two steps, especially in the vicinity of turning points, although its gains diminish as the quality of the indicators increases. Additionally, we also obtain decreasing returns of adding more indicators with similar signal-to-noise ratios. Using the four constituent series of the Stock-Watson coincident index, we illustrate these results for US data.

Keywords: Business Cycles, Output Growth, Time Series.

JEL Classification: E32, C22, E27.

1 Introduction

Burns and Mitchell's (1946) seminal work pointed out two business cycle characteristics: co-movements and the existence of two business cycle phases. To capture the first feature of the business cycle, Stock and Watson (SW, 1991), proposed a single-index linear dynamic factor model for industrial production, employment, income and sales. These four series have a common element that can be modeled by an underlying unobserved variable representing the overall economic activity, as in the Composite Index of Coincident Economic Indicators (CEI) of the Bureau of Economic Analysis (currently published by the Conference Board). A recent extension of this approach is the Aruoba, Diebold and Scotti (ADS, 2009) index of business conditions and the Chicago Fed National Activity Index (CFNAI).[1] To capture the second feature of the business cycle, Hamilton (1989) proposed a two-state Markov switching process, whose dynamics captures the two separate business cycle phases. Recent extensions of the model can be found in Hamilton (2011).

To capture both comovements and business cycle asymmetries, two alternative approaches have been used in the literature. The first approach was proposed in Diebold and Rudebusch (1996) and consists of a *two-step* estimation procedure. The first step is based on computing a coincident indicator by applying a linear factor model to a set of coincident indicators. In the second step, univariate Markov-switching techniques are applied to the coincident indicator. Examples of recent applications of Markov-switching techniques to linear factors are Diebold and Rudebusch (1996) in the case of SW; Brave and Butters (2012) in the case of a high frequency index such as ADS; Davig (2008) in the case of CFNAI; and Paap, Segers, and van Dijk (2009) in the case of CEI.

The second approach, which was initially proposed by Kim and Yoo (1995), Chauvet (1998) and Kim and Nelson (1998), is based on the natural extension of full dynamic-factor/Markov-switching models, which are estimated in *one step*. In their Markov-Switching Dynamic Factor Model (MS-DFM), comovements and business cycles are modeled with a nonlinear dynamic factor model whose common component is governed by an unobservable regime-switching variable that controls the business cycle dynamics. Recently, Chauvet and Hamilton (2006) and Chauvet and Piger (2008) examined the empirical reliability of these models in computing real-time inferences of the US business cycle.

---

[1] The Euro-STING model of Camacho and Perez Quiros (2010) is an European extension of these models.

Although inferring recession probabilities from the one-step approach is conceptually appealing, it is very tempting to fit a linear DFM to the economic indicators and a univariate Markov-switching model to the resulting linear coincident indicator. One reason is that the linear coincident indicators, such as ADS, CFNAI and CEI, are already constructed by several agencies and it seems straightforward to use univariate Markov-switching filters to compute state probabilities from them. Another reason is that the numerical algorithms used to evaluate the likelihood functions of the (one-step) MS-DFM usually suffer from the curse of dimensionality. If the analyst wants to perform business cycle analysis within a data rich environment (where the number of indicators can be above 100), the two-step procedure might be the only feasible procedure.

In spite of the coexistence of the two approaches, the analysis of their relative performance has still not been addressed. To fill in this gap, we examine the sources of misspecification of applying Markov-switching models to the common factor of a linear DFM when the data generating process is a nonlinear MS-DFM. For a given number of series, the two-step procedure faces greater difficulties to infer business cycle probabilities when the Kalman filter used to compute the linear factor model in the first step assigns large weights to past observations. Using the Riccati equation for the misspecified linear Kalman filter, we show that this typically occurs when the indicators are noisy. However, when the economic indicators are carefully selected to have large signal-to-noise ratios, the empirical performance of the one-step procedure is not expected to be highly superior to that of the two-step method. These theoretical results are confirmed by means of a Monte Carlo experiment.

We also examine the extent to which inferences about the state of the economy from nonlinear MS-DFM can be improved upon by including additional variables. In a linear framework, the recent literature provides mixed evidence. Boivin and Ng (2006) were among the first to show that the empirical forecasting performance of factor models does not necessarily improve with the number of indicators. In the nonlinear context of this paper, we show that the precision of the business cycle inference is expected to grow by including additional indicators, although with diminishing returns. However, regarding the quality of the new indicators, the expected gains from the additional indicators are progressively lower as the number of indicators already included increases. In empirical applications, this result implies that the gains from using large sets of indicators can be low compared to the computational complexity of dealing with nonlinear models that use many indicators.

Finally, we evaluate the empirical relevance of our results by using the four coincident economic indicators used in Stock and Watson (1991). Our main results are the following. First, as expected, since the indicators exhibit very high signal-to-noise ratios, the dynamics of the coincident indicators estimated from the two procedures are both in close agreement. Second, the two methods exhibit similar mean squared error measures when comparing the recession probabilities with a dummy that takes the value of one in the NBER recessions. Interestingly, when the analysis is restricted to the first month of the new state after a business cycle phase shift, the one-step model achieves mean squared errors reductions of about 20% (about 30% in the first month of each recovery). Third, we show the diminishing returns of adding more variables. Fourth, the results are confirmed in a real-time exercise, where the relative gains of the one-step estimation procedure are even larger.

The structure of this paper is organized as follows. Section 2 describes the sources of misspecification of a two-step estimation procedure when the data generating process is a nonlinear MS-DFM. In addition, it examines the extent to which the model performance could be improved by using additional indicators. Section 3 proposes a Monte Carlo experiment to analyze the ability of the models in computing business cycle inferences. Section 4 shows the empirical analysis of US business cycles. Section 5 concludes.

2 Two-step versus one-step approaches

This section examines the performance of two-step versus one-step estimation procedures of dynamic factor models with Markov-switching to accurately detect the probability of a given business cycle phase.[2]

2.1 Sources of misspecification

This section examines the sources of misspecification of the two-step estimation method when the procedure is used to infer the probability of a given business cycle phase. For this purpose, let us assume that the data generating process is a nonlinear MS-DFM but that an analyst erroneously fits a linear single-index DFM and tries to infer the recession probabilities from the resulting common factor with univariate Markov-switching techniques.

---

[2]The Appendix that shows all the expressions of this section is available at http://www.um.es/econometria/Maximo/.

Let $\mathbf{y}_t = (y_{1,t}, ..., y_{N,t})'$ be the vector of N economic indicators, which admits a single-index factor decomposition into a non-observed common factor f_t and N specific or idiosyncratic components:

$$
\begin{array}{ccccc}
\mathbf{y}_t & = & \mathbf{\Lambda} & f_t & + & \mathbf{u}_t \\
N \times 1 & & N \times 1 & 1 \times 1 & & N \times 1
\end{array}, \tag{1}
$$

where $\mathbf{\Lambda} = (\lambda_1, \lambda_2, ..., \lambda_N)'$ is the vector of factor loadings. Let $\mathbf{u}_t = (u_{1,t}, ..., u_{N,t})' \sim N(\mathbf{0}, \mathbf{\Sigma}_u)$ be multivariate Gaussian white noise. As in classical factor analysis, $\mathbf{\Sigma}_u = diag(\sigma_1^2, \sigma_2^2, ..., \sigma_N^2)$ is assumed to be a diagonal matrix.

To complete the specification of the data generating process, the factor is assumed to be governed by an unobserved regime-switching mean plus a noise

$$
f_t = \mu_{s_t} + a_t, \tag{2}
$$

where $a_t \sim N(0, \sigma_a^2)$ is an univariate Gaussian white noise. Within this framework, one can label $s_t = 0$ as expansions and $s_t = 1$ as recessions at time t if $\mu_0 > \mu_1$. The common dynamics of the coincident economic indicators are expected to exhibit high (usually positive) growth rates in expansions and low (usually negative) growth rates in recessions. In addition, s_t is assumed to evolve according to an irreducible 2-state Markov chain whose transition probabilities are defined by

$$
p(s_t = j | s_{t-1} = i, s_{t-2} = h, ..., I_{t-1}) = p(s_t = j | s_{t-1} = i) = p_{ij}, \tag{3}
$$

where $i, j = 0, 1$, and I_t is the information set up to period t.

The noises from equations (1) and (2) are uncorrelated for all leads and lags, that is, for all $i = 1, ..., N$, $E(u_{i,t}a_\tau) = 0, \forall t, \tau$. This assumption is standard and necessary to separate the common factors from the idiosyncratic noise, both in linear (see, among others, Stock and Watson, 1991, Diebold and Rudebusch, 1996, Aruoba, Diebold and Scotti, 2009) and in non-linear settings (see Kim and Yoo, 1995, Chauvet, 1998, and Kim and Nelson, 1998, among others).[3] The condition allows to decompose the variance of the observed series as the sum of two matrices: a rank deficient matrix, which is associated to the common factors, and a full rank matrix, which refers to the idiosyncratic noises. Typically, the variance of the idiosyncratic noises is either diagonal in small-N factor models or allows for for limited (weak) cross-correlation in large-N factor models.[4]

---

[3]Large factor models use the same identification condition (for instance, see Bai and Ng, 2013).

[4]See Heaton and Solo (2004) for identification of dynamic factor models.

Instead of fitting the one-step MS-DFM described above, let us assume that the analyst erroneously applies a two-step procedure as follows. In the first step, the analyst estimates a linear DFM to the set of N economic indicators whose common factor is assumed to follow a simple autoregressive process, which is assumed of order one to facilitate the analysis. Accordingly, the analyst computes a misspecified common factor, which is denoted with an asterisk,

$$f_t^* = d + \phi f_{t-1}^* + a_t^*. \tag{4}$$

In this expression, a_t^* is a univariate white noise with zero mean and variance $\sigma_{a*}^2 = 1$, which agrees with the standard identification assumption. The intercept d is added to take into account the possibility of a non-zero unconditional mean. The autoregressive parameter, ϕ, captures the serial correlation induced by the switching mean of the common factor described in (2).

Once $f_{t|t}^*$ is estimated in the first step, the analyst applies a univariate Markov-switching model to the common factor in the second step. Hence, the analyst estimates the nonlinear model by approximate maximum likelihood techniques and obtains the filtered state probabilities, which are computed to extract informative insights about the business cycle. Following Hamilton (1989), these probabilities can be expressed as

$$prob(s_t = j|I_t^*) = \frac{f(f_{t|t}^*|s_t = j, I_{t-1}^*)prob(s_t = j|I_{t-1}^*)}{f(f_{t|t}^*|I_{t-1}^*)}, \tag{5}$$

where $I_t^* = \left\{ (f_{\tau|\tau}^*)_{\tau=1}^{\tau=t} \right\}$, $f(f_{t|t}^*|s_t = j, I_{t-1}^*)$ is the conditional Gaussian density function

$$f(f_{t|t}^*|s_t = j, I_{t-1}^*) = \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp\left\{ -\frac{1}{2}\left(\frac{f_{t|t}^* - \mu_j}{\sigma_a^2}\right)^2 \right\}, \qquad j = 0, 1 \tag{6}$$

and $f(f_{t|t}^*|I_{t-1}^*)$ is the mixture

$$f(f_{t|t}^*|I_{t-1}^*) = \sum_{j=0}^{1} f(f_{t|t}^*|s_t = j, I_{t-1}^*)prob(s_t = j|I_{t-1}^*). \tag{7}$$

Therefore, the two-step estimation procedure faces a misspecification problem which might be of potential considerable importance in the detection of business cycle turning points.[5] Since $f_{t|t}^*$ is a linear combination of past and present values of $\mathbf{y}_t$, the density functions used in the univariate Markov-switching model depend not only on the current state but on all past and present states.

---

[5] To facilitate the analysis, we assume that population parameters are known and that the only source of misspecification comes from the way the common factor is extracted. Our simulations confirm that this is a very reasonable assumption.

Accordingly, the simple two-step method might underweight the signals of imminent changes in business cycle phases, implying longer delays in signaling business cycle phase changes.

To understand this statement, recall that the filtered linear common factor can be expressed as a weighted sum of past and present observations

$$f_{t|t}^* = d_0 + \sum_{\tau=1}^{t} \mathbf{w}_{t,\tau} \mathbf{y}_\tau, \tag{8}$$

where d_0 is included in the case that $d \neq 0$ and the weights $\mathbf{w}_{t,\tau}$ are the N-dimensional row vectors given by

$$
\begin{aligned}
\mathbf{w}_{t,t} &= \frac{1}{c_t}\mathbf{\Lambda}'\mathbf{\Sigma}_u^{-1}, \\
\mathbf{w}_{t,\tau} &= \frac{1}{c_\tau}\frac{1}{V_{\tau|\tau-1}}\phi\mathbf{w}_{t,\tau+1} = B_\tau\phi\mathbf{w}_{t,\tau+1},
\end{aligned}
\tag{9}
$$

for $\tau = t-1, ..., 1$. In the last expression, $V_{\tau|\tau-1}$ is the mean squared error of the misspecified state at τ with information up to $\tau-1$, $c_\tau = \frac{1}{V_{\tau|\tau-1}} + \mathbf{\Lambda}'\mathbf{\Sigma}_u^{-1}\mathbf{\Lambda}$, and $B_\tau = \frac{1}{c_\tau}\frac{1}{V_{\tau|\tau-1}}$.

The misspecification problem of the one-step estimation procedure is expected to be high when the weights decay slowly. In these cases, the approximation of the univariate Markov-switching model applied to the linear common factor will lead to important reductions in the timely detection of current turning points since the information of a potential regime switch contained in $\mathbf{y}_t$ diminishes its impact on $f_{t|t}^*$. Hence, a good strategy to examine the sources of misspecification of the two-step estimation procedure is to analyze the sources of persistence of the linear common factor which, according to expression (9), depends on B_τ and ϕ.

Focusing on B_τ, the larger it is, the stronger the misspecification of the two-step procedure becomes. To find the coefficients governing B_τ, it is worth noting that the linear filter reaches its steady state and $V_{t|t-1} = V$ since $\phi < 1$. Hence, the solution of the algebraic Riccati equation for the misspecified filter is

$$V = \frac{\sum_{i=1}^{N}\frac{\lambda_i^2}{\sigma_i^2} - (1-\phi^2) + \sqrt{\left(\sum_{i=1}^{N}\frac{\lambda_i^2}{\sigma_i^2} - (1-\phi^2)\right)^2 + 4\sum_{i=1}^{N}\frac{\lambda_i^2}{\sigma_i^2}}}{2\sum_{i=1}^{N}\frac{\lambda_i^2}{\sigma_i^2}}, \tag{10}$$

which implies that B_τ is

$$B_\tau = \frac{2}{\sum_{i=1}^{N}\frac{\lambda_i^2}{\sigma_i^2} + (1+\phi^2) + \sqrt{\left(\sum_{i=1}^{N}\frac{\lambda_i^2}{\sigma_i^2} - (1-\phi^2)\right)^2 + 4\sum_{i=1}^{N}\frac{\lambda_i^2}{\sigma_i^2})}}. \tag{11}$$

This expression reveals that the greater the autoregressive parameter ϕ and the sum $\sum_{i=1}^{N} \frac{\lambda_i^2}{\sigma_i^2}$, the smaller B_τ. To interpret the ratios $\frac{\lambda_i^2}{\sigma_i^2}$, it is worth writing the common factor as

$$f_{t|t}^* = \frac{1}{c_t} \left(\frac{1}{V_{t|t-1}} f_{t|t-1}^* + \sum_{i=1}^{N} \frac{\lambda_i^2}{\sigma_i^2} \left(\frac{y_{i,t}}{\lambda_i} \right) \right), \tag{12}$$

with $c_t = \frac{1}{V_{t|t-1}} + \sum_{i=1}^{N} \frac{\lambda_i^2}{\sigma_i^2}$. This expression states that $f_{t|t}^*$ is the weighted sum of two components. The first component, $f_{t|t-1}^*$, is the estimation of the factor at time t with information up to time $t-1$, and it has a weight which is proportional to the precision of this estimation. The second component is the weighted sum of the new information incorporated by the indicators observed at t, $\frac{y_{i,t}}{\lambda_i}$.[6] The weights , $\frac{\lambda_i^2}{\sigma_i^2}$, $i = 1, ..., N$, are called signal-to-noise ratios since they measure the precision of the indicators (inverse of the conditional variances $var(f_t|y_{i,t})$, $i = 1, ..., N$). Hence, the signal-to-noise ratios assign more weight to compute $f_{t|t}^*$ to the less noisy economic indicators. Then, when the signal-to-noise ratios are low, the weights to the past information, $\mathbf{w}_{t,\tau}$, become large. Summing up, using economic indicators with low signal-to-noise ratio leads the approximation of univariate Markov-switching dynamics of the common factor extracted from a linear DFM estimation to become increasingly more inappropriate to detect business cycle turning points in advance.

The influence of ϕ on the weights is twofold. According to expression (9), this parameter increases past weights directly. At the same time, it reduces the past weights through B_τ. Hence, the net effect of ϕ on business cycle identification is complex and we refer the readers to the simulation experiment developed in Section 3 for a more detailed analysis. In spite of this unclear effect, Timmermann (2000) showed that

$$\phi = \frac{(\mu_0 - \mu_1)^2 \, \pi_1 \pi_0 (p_{00} + p_{11} - 1)}{(\mu_0 - \mu_1)^2 \, \pi_1 \pi_0 + \sigma_a^2}, \tag{13}$$

where π_i is the steady state probability of state i, such that $\pi_1 + \pi_0 = 1$, and

$$\pi_i = \frac{1 - p_{jj}}{2 - p_{ii} - p_{jj}}, \tag{14}$$

with $i, j = 0, 1$. Accordingly, the autoregressive parameter is an increasing function of the difference of the within-state means and the persistence of the business cycle states.[7] Assuming

---

[6] Notice that $\frac{y_{i,t}}{\lambda_i}$ is the conditional expectation of the common factor from the i-th indicator $E(f_t|y_{i,t})$.

[7] It can be easily checked that for given probabilities, $p_{ii}, i = 0, 1$, the derivative of ϕ with respect to $(\mu_0 - \mu_1)^2$ is always positive as long as $p_{11} + p_{00} \geq 1$. This condition usually holds in empirical applications. In our empirical analysis (Section 4), we obtained $p_{00} = 0.98$ and $p_{11} = 0.89$.

that $\mu_0 > \mu_1$, this means that the larger the difference between the two conditional means (i.e., $\mu_0 >> \mu_1$), the larger ϕ should be and this could help to identify the business cycle regimes since it separates the Gaussians of the mixture (see Chauvet and Hamilton, 2006).

The effects of relaxing the assumptions about the dynamics of the idiosyncratic components, $\mathbf{u}_t$, deserve a final comment. In expression (1), we assumed that the idiosyncratic components followed a multivariate white noise. However, this assumption can be relaxed by appropriately defining $\mathbf{\Sigma}_u$ in the previous expressions. For instance, let us assume that the idiosyncratic components follow the VAR(p) process with $p = 1$

$$\mathbf{u}_t = \mathbf{\Psi}\mathbf{u}_t + \boldsymbol{\epsilon}_t, \tag{15}$$

where $var(\boldsymbol{\epsilon}_t) = diag(\sigma_1^2, ..., \sigma_N^2)$, and $\mathbf{\Psi} = diag(\psi_1, ..., \psi_N)$. In this case, $\mathbf{\Sigma}_u = diag(\frac{\sigma_1^2}{1-\psi_1^2}, ..., \frac{\sigma_N^2}{1-\psi_N^2})$. Notice that in this case, the conditional densities of the observed series depend on the hidden state both contemporaneously and through its first lag. We examine the case of idiosyncratic autoregressive noises in the simulations presented in Section 3.

2.2 The role of N

Empirical applications of MS-DFM frequently exhibit the typical curse of dimensionality problems of nonlinear models. This precludes the analysts from considering the case of large values of N. In spite of this comment, the question of how many economic indicators are useful to compute accurate inferences of business cycle turning points still holds. If an analyst starts with a set of $N - 1$ economic indicators that provides reasonable turning point signals, the problem reduces to the question of under which circumstances the additional N-th variable may be incorporated into the model leaving the dimension of the resulting nonlinear model manageable.

For this purpose, we consider that the set of N indicators is preferred to the set of $N - 1$ indicators if the former sufficiently increases the ability to appropriately detect true turning points and reduces the rate of false signals. Let us denote the set of $N - 1$ indicators by $I_{1,t} = \{y_{1,1}, ..., y_{1,t}; y_{2,1}, ..., y_{2,t}; ...; y_{N-1,1}, ..., y_{N-1,t}\}$, and the N-th indicator by $I_{2,t} = \{y_{N,1}, ..., y_{N,t}\}$. Hence, we find it useful to include the last indicator in inferring the state of the economy at time t whenever $prob(s_t = 1|I_{1,t}, I_{2,t}) > prob(s_t = 1|I_{1,t})$ when $s_t = 1$ (for instance, recessions) and $prob(s_t = 1|I_{1,t}, I_{2,t}) < prob(s_t = 1|I_{1,t})$ when $s_t = 0$ (for instance, expansions). Since it is straightforward to show that if $prob(s_t = i|I_{1,t}) = 1$, then $prob(s_t = i|I_{1,t}) = prob(s_t = i|I_{1,t}, I_{2,t}) = 1$, let us assume that $0 < prob(s_t = i|I_{1,t}) < 1$.

To start with, let us consider that the quality of the N-th indicator is similar to that of the set that contains the $N-1$ first indicators, i.e., that it is not a noisier time series. Let us focus the analysis on the identification of a given regime, for example on $s_t = 1$.[8] According to the Markov-chain properties of the model described in (1) to (3), $prob(s_t = 1 | I_{1,t}, I_{2,t}) = prob(s_t = 1 | y_{1,t}, ..., y_{N,t})$ and $prob(s_t = 1 | I_{1,t}) = prob(s_t = 1 | y_{1,t}, ..., y_{N-1,t})$. Then, the inference computed from the set of N indicators can be expressed as

$$
\begin{aligned}
prob(s_t &= 1 | I_{1,t}, I_{2,t}) = \frac{prob(s_t = 1, y_{N,t} | I_{1,t})}{f(y_{N,t} | I_{1,t})} \\
&= \frac{f(y_{N,t} | s_t = 1, I_{1,t})}{f(y_{N,t} | I_{1,t})} prob(s_t = 1 | I_{1,t}) \\
&= w_t prob(s_t = 1 | I_{1,t}),
\end{aligned} \tag{16}
$$

where $w_t = \frac{f(y_{N,t} | s_t = 1, I_{1,t})}{f(y_{N,t} | I_{1,t})}$. Therefore, the information of the new indicator $y_{N,t}$ will be useful to compute inferences about the business cycle at time t if $w_t > 1$ given that the true regime is $s_t = 1$.

It is useful to express w_t as

$$
w_t = \frac{f(y_{N,t} | s_t = 1, I_{1,t})}{f(y_{N,t} | s_t = 1, I_{1,t}) prob(s_t = 1 | I_{1,t}) + f(y_{N,t} | s_t = 0, I_{1,t})(1 - prob(s_t = 1 | I_{1,t}))}. \tag{17}
$$

Then, there exists informational content in the N-th indicator if

$$
f(y_{N,t} | s_t = 1, I_{1,t}) > f(y_{N,t} | s_t = 1, I_{1,t}) prob(s_t = 1 | I_{1,t}) + f(y_{N,t} | s_t = 0, I_{1,t})(1 - prob(s_t = 1 | I_{1,t})). \tag{18}
$$

This occurs whenever

$$
\frac{f(y_{N,t} | s_t = 1, I_{1,t})}{f(y_{N,t} | s_t = 0, I_{1,t})} > 1. \tag{19}
$$

Since the inequality in (19) does not necessarily hold for all possible values of $y_{N,t}$ (for instance, in the case of overlapping density functions) it will be useful to evaluate if this condition holds on average. Taking natural logarithms of this expression, the set of N indicators outperforms on average the set of $N-1$ indicators if

$$
\ln f(y_{N,t} | s_t = 1, I_{1,t}) - \ln f(y_{N,t} | s_t = 0, I_{1,t}) > 0 \tag{20}
$$

when $s_t = 1$. Under conditional Gaussianity

$$
f(y_{N,t} | s_t = i, I_{1,t}) = \frac{1}{\sqrt{2\pi\sigma_{N|1}^2}} \exp\left(-\frac{1}{2\sigma_{N|1}^2}\left(y_{N,t} - y_{N,t|t}^{(i)}\right)^2\right), i = 0, 1
$$

_____

[8]The treatment of regime $s_t = 0$ is symmetric.

where $y_{N,t|t}^{(i)}$ and $\sigma_{N|1}^2$ are the mean and variance of the conditional density function, given in the Electronic Appendix. Taking into account all possible outcomes of $y_{N,t}$ when $s_t = 1$, the expected value of the difference between the two conditional densities under conditional Gaussianity is given by

$$
\int \ln f(y_{N,t}|s_t = 1, I_{1,t}) f(y_{N,t}|s_t = 1, I_{1,t}) dy_{N,t} \tag{21}
$$
$$
- \int \ln f(y_{N,t}|s_t = 0, I_{1,t}) f(y_{N,t}|s_t = 1, I_{1,t}) dy_{N,t} > 0.
$$

Therefore, one could evaluate if condition (19) is fulfilled on average.

The next proposition, quantifies the expected gains in terms of business cycle identification of adding a new economic indicator $y_{N,t}$ when $s_t = 1$ to a given set of $N-1$ indicators. The magnitude of the change is a measure of the averaged informational content of $y_{N,t}$ about the two states (its ability to separate them) and uses the concept of conditional entropy (or Kullback-Leibler divergence).

Proposition 1 *Under conditional Gaussianity, the Kullback-Leibler (KL) divergence of $f(y_{N,t}|s_t = 0, I_{1,t})$ with respect to $f(y_{N,t}|s_t = 1, I_{1,t})$ under the MS-DFM assumptions described in (1) to (3), is given by*

$$
KL = \frac{\lambda_N^2 (\mu_0 - \mu_1)^2}{2\sigma_N^2} \frac{\frac{1}{\sigma_a^2}}{\frac{1}{\sigma_a^2} + \sum_{i=1}^{N-1} \frac{\lambda_i^2}{\sigma_i^2}} \frac{\frac{1}{\sigma_a^2}}{\frac{1}{\sigma_a^2} + \sum_{i=1}^{N} \frac{\lambda_i^2}{\sigma_i^2}}. \tag{22}
$$

This expression implies that (i) if there are separate business cycles regimes, in the sense that $\mu_1 \neq \mu_0$ and $\sigma_a^2 < \infty$; and (ii) if the new indicator is informative, in the sense that $\lambda_N \neq 0$ and $\sigma_N^2 < \infty$, then the divergence is strictly positive. This implies that adding a new indicator is (on average) always useful in terms of business cycle identification.

Interestingly, each additional indicator does not increase the Kullback-Leibler divergence in the same proportion. The informational content of the additional N-th indicator increases with the signal-to-noise ratio of this indicator. Hence, the accuracy of the model to provide clear business cycle signals increases with the quality of the new indicator. However, assuming that the signal-to-noise ratio is the same for all the economic indicators, the divergence is a decreasing function of the number of indicators. Therefore, the gains of adding new indicators can be lower than proportional to the number of indicators that have already been included to infer the business cycle probabilities. Hence, the decreasing ability to improve upon the accuracy of the model cannot always be compensated with the quality of the new indicator. This implies

that the gains from using large sets can be low compared with the computational complexity of handling with many indicators in empirical applications of nonlinear models.

Figure 1 helps us to assess the size of the gains as a function of both the signal-to-noise ratio of the new indicator and the number of indicators already included in the model. To facilitate comparisons we set $\mu_0 - \mu_1 = \sqrt{2}$ and $\sigma_a^2 = 1$, and we assume that the signal-to-noise indicators included in $I_{1,t}$ is 1 for all of its $N - 1$ indicators. Using these assumptions, the figure plots the KL divergence of $f(y_{N,t}|s_t = 0, I_{1,t})$ with respect to $f(y_{N,t}|s_t = 1, I_{1,t})$ when the additional indicator exhibits potential signal-to-noise ratios from 0.1 to 40. Given a number of indicators, for instance $N = 2$, the figure shows that the divergence increases rapidly when the new indicator exhibits signal-to-noise ratios of up to about six times the signal-to-noise ratios of the existing indicators. However, the divergence becomes hump-shaped at these values, implying that using less noisy indicators does not help to increase the divergence by a large amount.

To evaluate the role of the number of indicators, the figure also plots the results from repeating the previous exercise when number of indicators also increases up to $N = 5$ and $N = 8$. The figure shows that the divergence functions are shifted down by increasing N. This implies that although enlarging the original set of indicators with a new indicator of fixed signal-to-noise ratio leads improves the accuracy to infer the business cycle phases, the gains are low when the initial set of indicators becomes large.

These findings resemble those suggested by Poncela and Ruiz (2012) in a linear framework. They show that, when the parameters are known, the filter uncertainty is a non-increasing function of the cross-sectional dimension, although it only decreases marginally in many empirical situations when the number of time series becomes modestly large.

3   Monte Carlo simulations

In this section, we set up several Monte Carlo experiments to study how the data might affect the empirical performance of one-step versus two-step estimation procedures as well as the role of N to infer business cycle probabilities. For this purpose, we generate a total of $M = 1000$ sets of N idiosyncratic components $\mathbf{u}_t^m$ of length $T = 200$ and equal variances $\sigma_i^2 = \sigma^2$. The dynamics of these time series are assumed to follow autoregressive processes of order one with autoregressive parameters equal to 0.3. In addition, we generate $M = 1000$ dummy variables b_t^m of zeroes and ones of length $T = 100$, which are used to simulate different sequences of

expansions and recessions. To ensure that the dummies share the US business cycle properties, we assume that b_t^m follows Markov chains with $p_{00} = 0.9$ and $p_{11} = 0.7$.[9] Hence, we generate $M = 1000$ common factors that follow Markov-switching processes f_t^m by using the business cycle sequences b_t^m and by assuming that $\sigma_a^2 = 1$. 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 time series $\mathbf{y}_t^m$. Then, we apply both the two-step and one-step estimation procedures to extract the filtered probabilities of state 1, $p_{t,i}^m$, with $i = I, II$ in the cases of using one-step and two-step estimation procedures, respectively. The Monte Carlo experiment is developed for $N = 3$ indicators. We generate indicators of different differences of the within-state means ($\mu_0 - \mu_1 = 1, 2, 4$ and 10), and different variances ($\sigma^2 = 0.5, 1.5$ and 4.5).

To examine the relative accuracy of these approaches in situations of empirical interest, we assume that the model parameters are unknown and have to be estimated by maximum likelihood.[10] The model is written in state space form and the unobserved components are estimated by using the Kalman filter. Since each iteration of the filter produces a 2-fold increase in the number of cases to consider, we make the Kalman filter operable by following Kim (1994). Therefore, we collapse terms by computing a weighted average of the updating procedures by the probabilities of the Markov state, in which the mixture of four Gaussian densities is collapsed, after each observation, into a mixture of two densities.

For each m-th replica, we quantify the ability of these two estimation procedures to detect the actual state of the business cycle by computing the Quadratic Probability Score (QPS):

$$QPS_i = \frac{1}{M} \sum_{m=1}^{M} \frac{1}{T} \sum_{t=1}^{T} (p_{ti}^m - b_t^m)^2, \tag{23}$$

where $i = I$ for the one-step estimation procedure, and $i = II$ for the two-step estimation procedure. This measure can be interpreted as the average over the M replications of the squared deviation from the generated business cycles.

To examine the sources of misspecification of the two-step estimation method, Table 1 displays the QPS statistics, which are computed for the different scenarios described above. In addition, since turning point detection is a key business cycle question, the table shows in brack-

---

[9]Using the NBER dates, these figures are the percentage of quarters classified as expansions that are followed by expansions and the percentage of quarters classified as recessions that are followed by recessions in the period 1959.3-2010.3.

[10]We repeated the analysis by assuming that the parameters were known. For reasonable values of these parameters we obtained qualitative similar results, which have been omitted to save space.

ets the scores when QPS is calculated only for those t that refer to the first period after the phase shifts.

The main message of this table is that the one-step estimation procedure unequivocally performs better than the two-step procedure. However, the relative performance gains depend on the quality of the indicators used in the business cycle analysis. For a given idiosyncratic variance, higher differences of within-state means (from 1 to 10) improve the performance of both one-step and two-step estimation procedures. When these differences become large enough, although the one-step approach improves proportionally more than the two-step procedure, the gains are statistically important but not economically meaningful since the misspecified two-step approach is already very accurate. For example, for a variance of $\sigma^2 = 1.5$, the ratio of the QPS statistics between the two-step and the one-step procedures when $\mu_0 - \mu_1 = 1$ is 1.24 ($QPS_{II} = 0.250$ versus $QPS_I = 0.202$) while the ratio is more than 5000 when $\mu_0 - \mu_1 = 10$ ($QPS_{II} = 0.023$ versus $QPS_I = 4.22E\text{-}06$). However, the QPS of the two-step approach was very low already (0.023).[11] The results also hold for the turning point detection whose corresponding outcomes are displayed in brackets in Table 1. In this case, the table also shows that when the set of indicators included in the analysis are very precise in terms of signal-to-noise ratios, there is less room for the one-step method to improve the empirical performance of the two-step procedure. If the indicators are not so good, the empirical performance of the one-step procedure can significantly outperform that of the two-step procedure.

From Table 2, which examines the role of N in the one-step estimation procedure performance, there are two noteworthy findings that deserve comments. First, the table shows that the number of indicators used to infer the business cycle phases matters since increasing the number of time series leads to business cycle identification improvements, i.e., QPS reductions. However, the usefulness of new indicators in terms of business cycle identification refinements dramatically decreases when the number of indicators already used becomes large. For example, for $\sigma^2 = 1.5$, a model of $N = 1$ indicator exhibits $QPS_I^{N=1} = 0.168$, and adding two more indicators implies an improvement of 26% ($QPS_I^{N=3} = 0.124$). Nevertheless, the improvement falls to 12% when the set of indicators is extended from $N = 3$ to $N = 5$, and it is only 9% when the number of indicators is enlarged from $N = 5$ to $N = 7$. Second, the table highlights that the quality of the indicators matters when one is interested in quantifying the expected gain from

---

[11]To make the QPS results more readable, recall that the inference that gives probability of recession equals to zero for all t leads to a QPS of 0.20.

14

enlarging the model. In that sense, when the indicators are very precise (for instance, $\sigma^2 = 0.5$), the gain of enlarging the model from $N = 3$ to $N = 7$ indicators is only 9%. However, when the indicators are not very precise (for instance, $\sigma^2 = 4.5$), using $N = 7$ instead of $N = 3$ indicators increases the expected accuracy by more than 27%. The intuition for this result is that the model with $N = 3$ noiseless indicators is able to infer the business cycle with considerable accuracy and does not leave too much room for any improvement when the number of indicators is enlarged. In the case of computing business cycle inferences from noisier indicators, there are larger potential accuracy gains from a model with increased dimension.

4   Empirical results

The purpose of this section is to examine the empirical performance of the one-step versus the two-step estimation procedures and the role of combining information from a set of economic indicators by using an updated real-time version of the data set previously used by Stock and Watson (1991), Chauvet (1998) and Chauvet and Piger (2008). The indicators are monthly industrial production index (IP), nonfarm payroll employment (EMPL), real personal income less transfer payments (INC) and real manufacturing and trade sales (SALES) from 1967.01 to 2010.11.[12]

To start with, we fit a linear factor model to one hundred times the change in the natural logarithm of these four macroeconomic variables.[13] The maximum likelihood estimates, which are displayed in the top panel of Table 3, show that the signal-to-noise ratios are in line with the values used in our simulations. In particular, the signal-to-noise ratio of IP, EMPL, INC and SALES are 1.83, 0.65, 0.09, and 0.37, which would (roughly) correspond in the simulations to the cases of $\sigma^2 = 0.5$ (IP), $\sigma^2 = 1.5$ (EMPL) and $\sigma^2 = 4.5$ (INC and SALES) since they were computed with $\lambda = 1$. The fact that the magnitude of the parameters chosen in some of the simulations exercise matches the data reinforces the findings of the simulations exercise whose results can directly be applicable to the empirical analysis. The common factor obtained from the two-step approach is depicted in Figure 2 (left scale).

Then, we fit a Markov-switching model to the estimated factor. The maximum likelihood

---

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

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

estimates, which are also reported in the top panel of Table 3, show that the transition probabilities are very persistent ($p_{00} = 0.98$, $p_{11} = 0.89$) and that the within-state means are separate from each other ($\mu_0 = 0.32$, $\mu_1 = -1.78$). According to our simulation results, the high signal-to-noise ratios and the separate within-state means would help the two-step estimation procedure to compute accurate inferences of the US business cycle dates. Figure 2 also plots the probabilities that the coincident indicator is in the negative growth rate (right scale), which are in striking accord with the professional consensus as to the history of US business cycles.

In contrast to this two-step procedure, we also fit a multivariate dynamic factor model with regime switching in which the two key features of the business cycle are encompassed and estimated in one step. The maximum likelihood estimates of this model are reported in the bottom panel of Table 3 along with their standard errors. Figure 3 plots the nonlinear coincident indicator (left scale), which also tracks the business cycle well, with pronounced drops that synchronously correspond to the NBER-designated recessions. The filtered probabilities that the coincident indicator is in the negative growth rate (right scale) also show remarkable success in matching the NBER reference dates.

Although the filtered probabilities seem to indicate that both estimation procedures reproduce the NBER chronology very closely, Table 4 displays the QPS statistics, which evaluate formally the performance of the models.[14] The entries of the table reveal the high performance score of both models ($QPS_I = QPS_{II} = 0.05$), which implies that there is a similar correspondence between the probabilities inferred from the models and the business cycle realizations.

Since the previous results are computed as averages over the entire sample, the analysis may fail to identify the ability of the forecasting models to evaluate the odds of the occurrence of important events such as the turning point dates. To gain some insight into how quickly the business cycle turning points are identified, the probability scores are also computed on a restricted sample that only includes the first month of each expansion and the first month of each recession. The results are also displayed in Table 4 and show that the highest performance score is now achieved by the full Markov-switching dynamic-factor model, which assesses QPS reductions of about 20% ($QPS_{II} = 0.39$ versus $QPS_I = 0.31$).[15] These results are in accordance

---

[14]In the empirical analysis, QPS is defined as in the Monte Carlo analysis with $M = 1$.

[15]The improvements are even clearer when the analysis is restricted to the first month after the trough. In this case, the reductions are of almost 30% ($QPS_{II} = 0.50$ versus $QPS_I = 0.36$), although they are omitted from Table 4 to save space. This result could be connected with rapid growth in the recoveries documented by Sichel (1994), Kim and Nelson (1999), and Morley and Piger (2006).

with the theoretical analysis of Section 2 and the simulations of Section 3.[16]

As we stated in the theoretical and Monte Carlo analyses, we now address the role of the signal-to-noise ratio (SNR) and the number of indicators used to compute the business cycle inferences. To consider the role of the signal-to-noise ratios, we start by fitting a model that includes only the two indicators with intermediate SNR, i.e., SALES and EMP. In this case, the model achieves a QPS of 0.089. Enlarging the model with the indicator that exhibits the highest SNR, IP, implies QPS reductions of 38% ($QPS = 0.055$). However, enlarging the model with the indicator that shows the lowest SNR, INC, implies QPS reductions of only 25% ($QPS = 0.067$).

To illustrate the diminishing returns of adding more indicators, we examine the effect of enlarging different combinations of indicators with IP, the indicator with the largest SNR. When a model that uses EMP and INC is enlarged with IP, the QPS reduction is of 42% ($QPS = 0.096$ versus $QPS = 0.056$). However, when IP is added to a model that uses EMP, INC and SALES the reduction is of only 22% ($QPS = 0.067$ versus $QPS = 0.052$). In addition, when low SNR is combined with the diminishing returns to adding more variables, the usefulness of new indicators in terms of business cycle identification refinements dramatically decreases. For example, the model with the two indicators that exhibit the largest SNR, IP and EMP achieves a QPS of 0.091. Adding the indicator with the following SNR, SALES, achieves a QPS of 0.056, which implies a reduction of 38%. However, when the noisier indicator (INC) is added, the model that uses the four indicators reaches a QPS of 0.052, which implies an additional reduction of only 7%.

To address what might happen by including a lot more observables, we perform a two-step estimation of the large data base used by Stock and Watson (2012).[17] The sample runs from 1967.03 to 2011.06 and the series were processed as indicated by these authors. The first

---

[16]To verify the robustness of these results, we repeated the analysis for the five monthly time series of the Euro-STING model proposed in Camacho and Perez-Quiros (2010) that are available for the longer sample (1991.04 to 2014.03). The series are Economic Sentiment Indicator of the European Commission, Belgium overall business indicator, German IFO and growth rates of Industrial Production and Exports. Each series reaches a QPS of 0.19, 0.21, 0.20, 0.18 and 0.17, respectively. The one-step and two-step procedures provide $QPSs$ of 0.15 and 0.17, respectively.

[17]We thank a referee for this suggestion. The data were downloaded from

http://www.princeton.edu/~mwatson/publi.html

latent factor and the filtered probabilities of recession inferred from the factor are plotted in Figure 4. The QPS is 0.100, which doubles the QPS obtained from our four indicators.[18] Removing IP from this large data implies that QPS increases only 5% ($QPS = 0.105$) while removing the whole category of industrial production data implies that QPS increases only 15% ($QPS = 0.115$).

Finally, the real-time performance in tracking the US business cycle is now evaluated by using a real-time data set. That is, the inferences are computed monthly over the past 35 years by using only the data that would have been available at the month being considered. This is accomplished by estimating the models recursively with new data vintages and evaluating the evidence for a new turning point at the one-period ahead forecast in every period, following Chauvet and Piger (2008). This method provides a more realistic assessment of how the models would have performed, as it does not assume the knowledge of data revisions that were not available at the time the model would have been used. The real-time forecasts are computed for the period 1976.10-2010.11.

Figures 5 and 6 show that the real-time results are of the same nature as the in-sample results. Overall, the two-step and the one-step procedures exhibit similar forecasting accuracy. However, there is a reduction of about 5% in QPS when the analysis is restricted to examining the ability of the models to compute turning points inferences from the one-step method with respect to the two-step method.[19] Again, the overall relative real-time performance of multivariate models with respect to univariate models reveals that the former outperform the latter. According to Table 4, the QPS increase from 0.06 in the case of multivariate models to 0.08, 0.12, 0.16, and 0.09 in the case of univariate Markov-switching models of IP, EMPL, INC, and SALES, respectively. The real-time gains of combining information from several economic indicators are of the same order of magnitude as the in-sample gains.

5 Conclusions

We show that mixing information helps to compute business cycle inferences even in a real-time analysis. There are two ways of mixing this information. The first one is to compute

---

[18]According to Poncela and Ruiz (2012), the performance of the models can diminish when the number of parameters to be estimated increases dramatically.

[19]The reductions in QPS are of more than 40% ($QPS_{II} = 0.30$ versus $QPS_I = 0.16$, not included in Table 4) when the analysis is restricted to the periods right after the troughs.

Markov-switching probabilities from a coincident indicator, which is the outcome of a linear dynamic factor model. The second one is to compute the probabilities directly from a full Markov-switching dynamic-factor model. Although the first approach is computationally easier to implement, the second strategy is conceptually more appealing. In this paper, we examine the circumstances under which one is preferred to the other.

We find that the full Markov-switching dynamic-factor model exhibits higher business cycle performance, especially at the turning points. However, we show that the larger the quality of the business cycle indicators used in the analysis, the closer the ability of the two approaches to track the business cycles. This implies that, when the set of indicators included in the analysis are good indicators of the business cycle, the overall differences between the two approaches diminish considerably. Even in this case, the superior ability of the full Markov-switching dynamic-factor model appears in the turning points. In addition, we also show that the more variables included in the model, the better the fit. However, the improvements of adding new indicators in terms of business cycle identification dramatically decrease with the number of indicators already used, even in the case of incorporating indicators with high signal-to-noise ratios.

Using the four constituent indicators of the Stock and Watson (1991) coincident index, production, employment, income and sales, we examine the empirical performance of the models to assess the US business cycle. Since these indicators are of very high quality, we expected a high business cycle performance of the two methods. Consistent with this, we show that the business cycle dates obtained from their respective filtered probabilities are in both cases almost identical to the officially recognized US business cycle chronology. However, we obtain some relative gains of the full Markov-switching dynamic-factor model when the analysis is focused on turning points detection, one of the main goals of Markov-switching models. We confirm this result in a real-time analysis of the models' performance over the past 35 years.

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] Bai, J. and Ng, S. 2013. Principal Components Estimation and Identification of the Factors. *Journal of Econometrics* 176: 18–29.

[3] Boivin, J. and Ng, S. 2006. Are more data always better for factor analysis? *Journal of Econometrics* 132: 169-194.

[4] Brave, S., and Butters, R. 2012. Diagnosing the Financial System: Financial Conditions and Financial Stress. *International Journal of Central Banking*, 8(2), 191-239.

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

[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] Camacho, M., and Perez Quiros, G. 2010. Introducing the Euro-STING: Euro area Short Term Indicator of Growth *Journal of Applied Econometrics* 25: 663-694.

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

[9] 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.

[10] 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.

[11] Davig, T. 2008. Detecting recessions in the Great Moderation: a real-time analysis. *Federal Reserve Bank of Kansas City Economic Review*, Fourth Quarter: 5-33.

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

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

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

[15] Heaton, C. and Solo, V. 2004. Identification of causal factor models of stationary time series. *Econometrics Journal* 7: 618-627.

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

[17] 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.

[18] Kim, Ch., Nelson, Ch,. 1999. Friedman's plucking model of business fluctuations: Tests and estimates of permanent and transitory components. *Journal of Money, Credit, and Banking* 31: 317-334.

[19] Kim, M., 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] Morley J, Piger J. 2006. The Importance of Nonlinearity in Reproducing Business Cycle Features. Contributions to Economic Analysis Series, Vol. 276, Nonlinear Time Series Analysis of Business Cycles. Elsevier: Amsterdam and San Diego.

[21] Paap, R., Segers, R. and van Dijk, D. 2009. Do leading indicators lead peaks more than troughs? *Journal of Business and Economic Statistics* 27: 528-543.

[22] Poncela, P., and Ruiz, E. 2012. More is not always better: back to the Kalman filter in dynamic factor models. UC3M Working papers, Statistics and Econometrics, 12/17.

[23] Sichel, D., 1994. Inventories and the three phases of the business cycle. *Journal of Business and Economic Statistics* 12: 269-277.

[24] 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.

[25] Stock, J., and Watson, M. 2012. Disentangling the channels of the 2007-2009 recession, *Brookings Papers on Economic Activity*, Spring 2012.

[26] Timmermann, A 2000. Moments of Markov switching models. *Journal of Econometrics* 96: 75-111.

Table 1. Two-step versus one-step estimation procedures

σ^2	One step				Two steps			
	$\mu_0-\mu_1$							
	1	2	4	10	1	2	4	10
0.5	0.201	0.097	0.016	5.09E-08	0.229	0.099	0.045	0.018
	(0.332)	(0.255)	(0.070)	(1.29E-08)	(0.357)	(0.261)	(0.240)	(0.121)
1.5	0.202	0.124	0.027	4.22E-06	0.250	0.131	0.056	0.023
	(0.326)	(0.275)	(0.095)	(2.67E-05)	(0.363)	(0.292)	(0.272)	(0.146)
4.5	0.204	0.171	0.061	4.02E-04	0.279	0.201	0.087	0.031
	(0.320)	(0.307)	(0.169)	(1.47E-03)	(0.394)	(0.366)	(0.331)	(0.191)

Notes. Entries show the average over the replications of the averaged squared deviation of filtered probabilities of low-mean state from the 1000 generated business cycle sequences. The results when the analysis is restricted to the first month after phase shifts are in brackets. "One step" refers to MS-DFM and "two steps" refers to MS applied to a linear common factor. The replications use three indicators (N=3). Parameters are estimated.

Table 2. The role of N on MS-DFM

σ^2	N			
	1	3	5	7
0.5	0.122	0.097	0.092	0.088
1.5	0.168	0.124	0.109	0.099
4.5	0.202	0.171	0.141	0.124

Notes. Entries show the average over the replications of the averaged squared deviation of filtered probabilities of low-mean state from the 1000 generated business cycle sequences. N denotes the number of variables included in the model, and σ^2 the variance of the idiosyncratic shocks. The model has been generated with $\mu_0-\mu_1=2$. Parameters are estimated.

Table 3. Maximum likelihood estimates

Two-step procedure						
		Factor	IP	Empl	Inc	Sales
DMF	λ_i	-	0.69 (0.03)	0.50 (0.03)	0.28 (0.03)	0.45 (0.03)
	ϕ_1	0.43 (0.06)	-0.26 (0.07)	0.23 (0.03)	-0.20 (0.02)	-0.36 (0.04)
	ϕ_2	0.22 (0.05)	-0.21 (0.08)	0.53 (0.04)	-0.05 (0.04)	-0.15 (0.05)
	σ_i^2	1	0.26 (0.04)	0.24 (0.02)	0.85 (0.03)	0.59 (0.03)
MS		μ_1	μ_2	$\sigma_{a^*}^2$	p_{00}	p_{11}
		0.32 (0.04)	-1.78 (0.14)	0.80 (0.05)	0.98 (0.01)	0.89 (0.04)

One-step procedure					
		IP	Empl	Inc	Sales
Indicators	λ_i	0.69 (0.03)	0.42 (0.02)	0.28 (0.04)	0.46 (0.03)
	ϕ_1	-0.18 (0.08)	0.24 (0.03)	-0.20 (0.02)	-0.34 (0.04)
	ϕ_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_1	μ_2	$\sigma_{a^*}^2$	p_{00}	p_{11}
	0.32 (0.07)	-2.00 (0.20)	1	0.98 (0.01)	0.85 (0.05)

Note: Standard errors are in brackets.

Table 4. Empirical performance

	1-step	2-steps	IP	Empl	Inc	Sales
In sample (1967.01-2010.11)						
Total	0.05	0.05	0.07	0.12	0.13	0.09
Turning points	0.31	0.39	0.54	0.67	0.42	0.33
Real time (1976.10-2010.11)						
Total	0.06	0.06	0.08	0.12	0.16	0.09
Turning points	0.36	0.54	0.49	0.73	0.39	0.36

Note. Entries labeled as "total" refer to QPS statistics. In the case of entries labeled as "turning points", the QPS is computed by using the first month after the phase shifts.

Figure 1. KL divergence
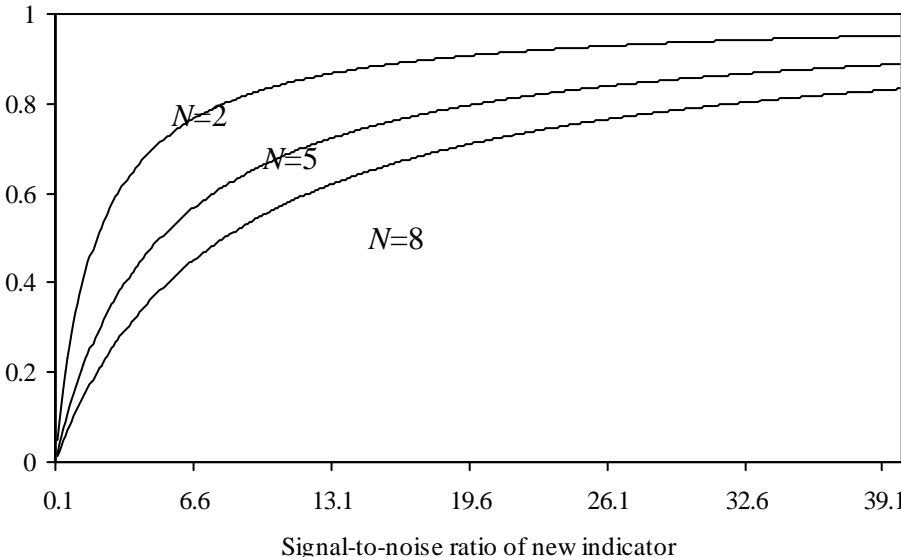
Note. This figure measures the gains of adding a new indicator (the initial set contains *N*-1 indicators) to infer the probability of recession when a recession occurs as a function of its signal-to-noise ratio.
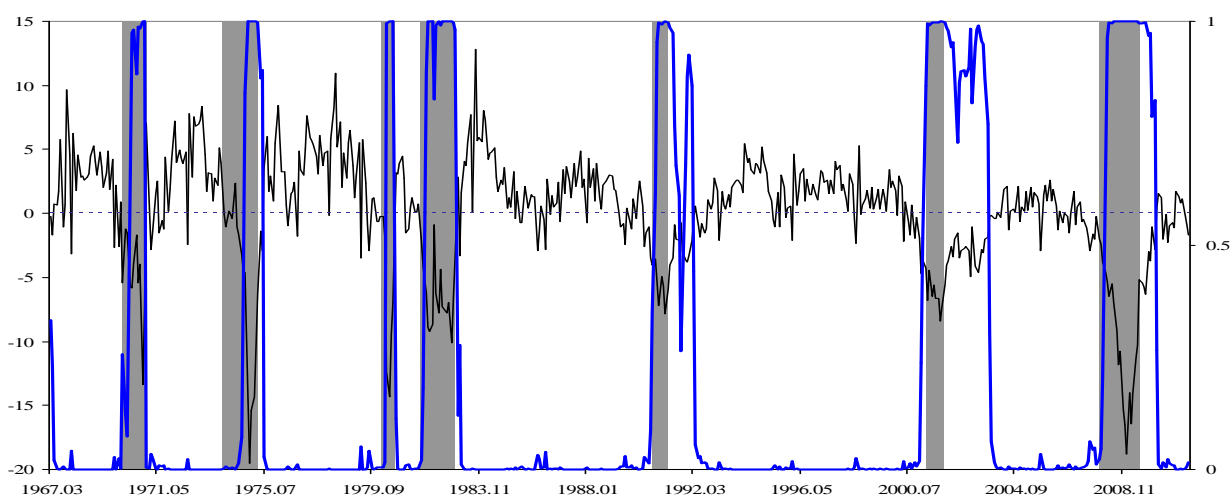
Figure 2. Two-step method
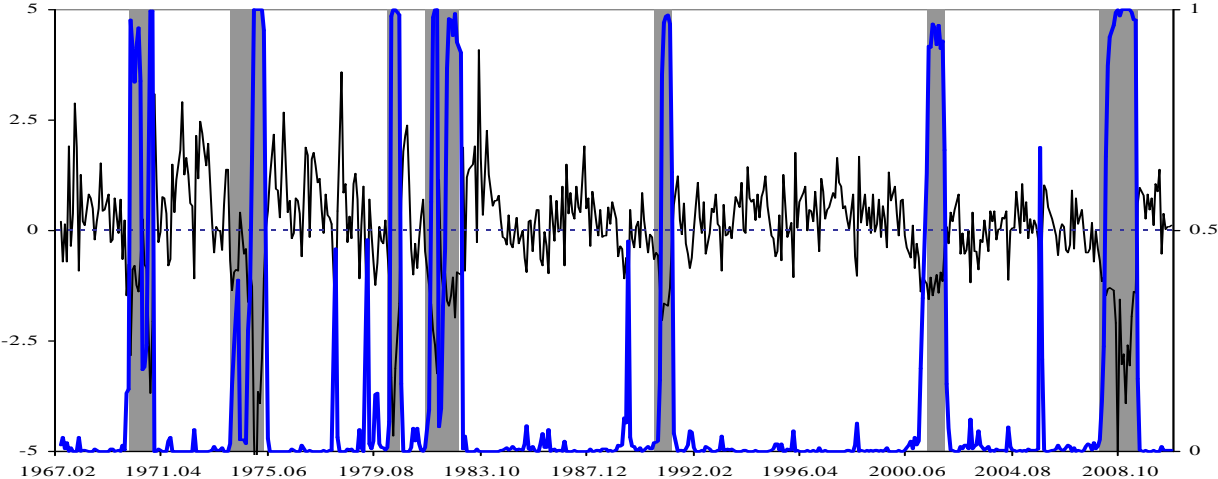
Notes: This graph plots the common factor (left scale) and the filtered probabilities of recession (right scale) computed from the two-step procedure. Shaded areas correspond to recessions as documented by the NBER.

Figure 3. One-step method
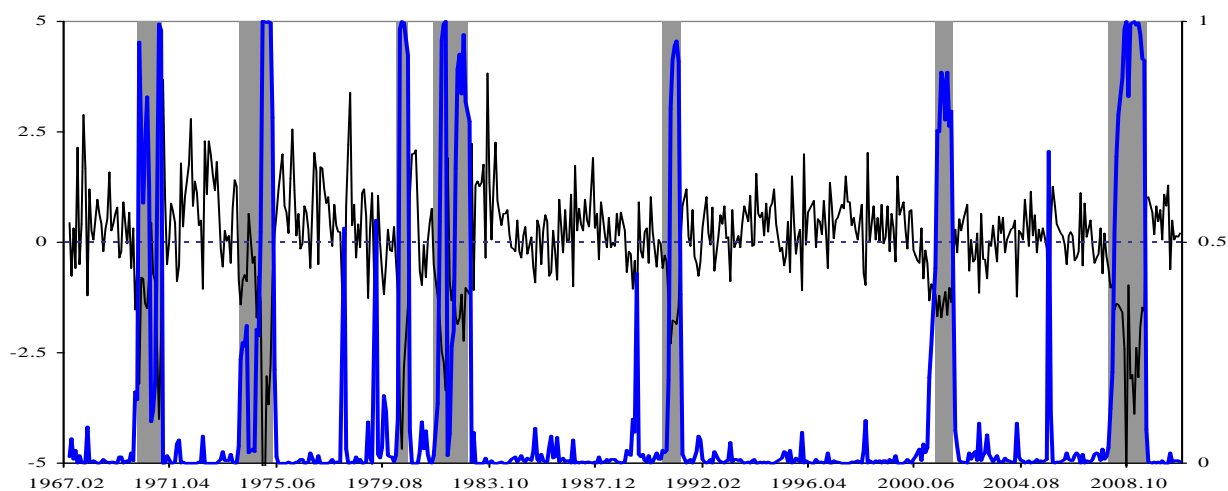
Notes: This graph plots the common factor (left scale) and the filtered probabilities of recession (right scale) computed from the one-step procedure. Shaded areas correspond to recessions as documented by the NBER.

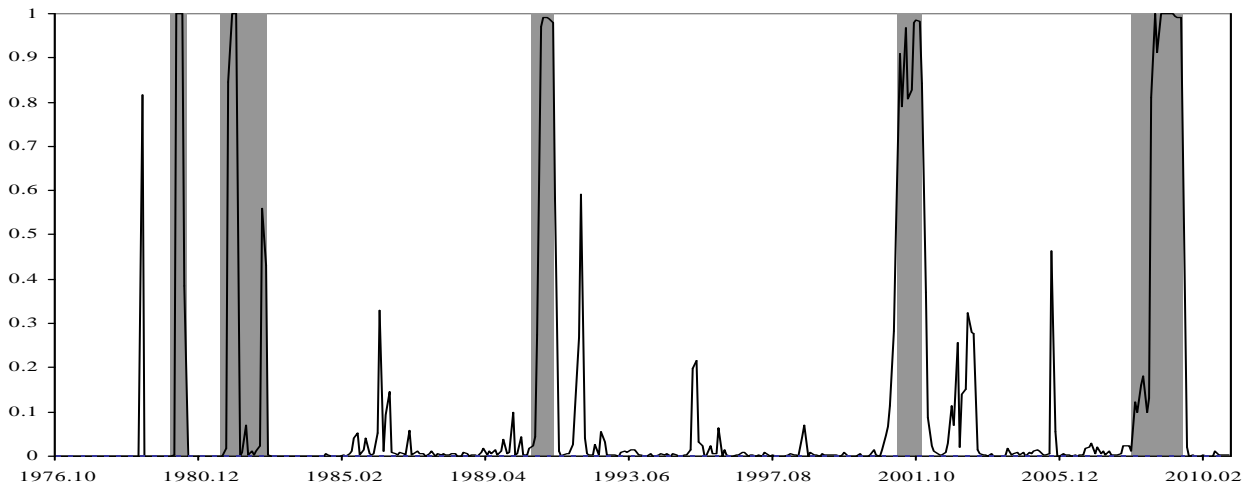Figure 4. Stock and Watson (2012) data base

Notes: This graph plots the first factor (left scale) and the filtered probabilities of recession (right scale). Shaded areas correspond to recessions as documented by the NBER.

Figure 5. One-period-ahead real-time forecasts from one-step model
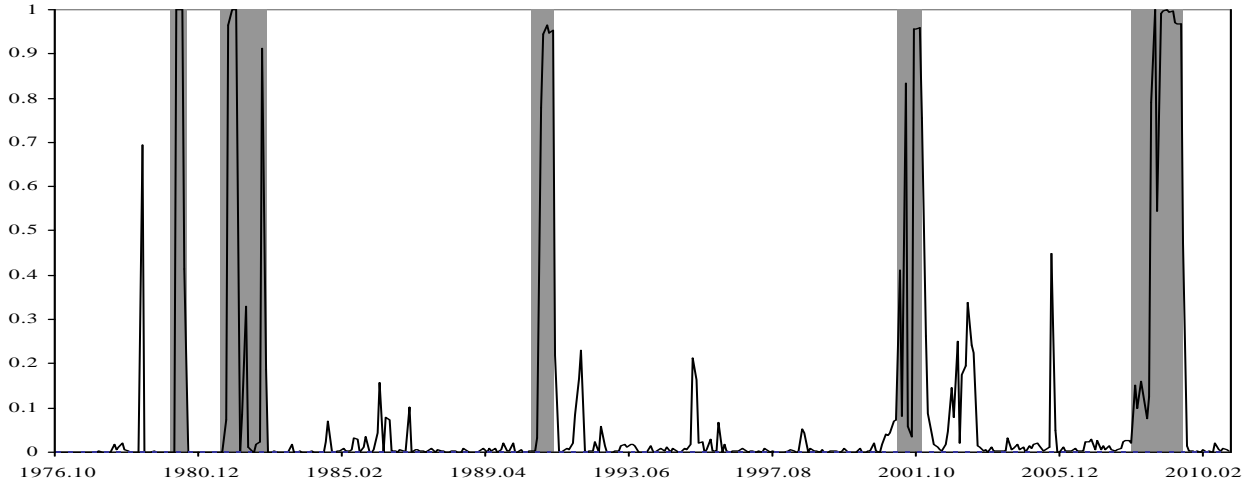
Note: The graph shows the one-period ahead real-time forecasts of the probability of being in recession. Shaded areas correspond to recessions as documented by the NBER.

Figure 6. One-period-ahead real-time forecasts from two-step model

Note: The graph shows the one-period ahead real-time forecasts of the probability of being in recession. Shaded areas correspond to recessions as documented by the NBER.