ARTICLE IN PRESS

International Journal of Forecasting xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Accelerating peak dating in a dynamic factor Markov-switching model☆

Bram van Os *, Dick van Dijk

Econometric Institute, Erasmus University Rotterdam, Netherlands

ARTICLE INFO

Keywords:
Business cycles
Generalized autoregressive score models
Time-varying transition probabilities
Turning points
Term spread

ABSTRACT

The dynamic factor Markov-switching (DFMS) model introduced by Diebold and Rudebusch (1996) has proven to be a powerful framework for measuring the business cycle. We extend the DFMS model by allowing for time-varying transition probabilities, intending to accelerate the real-time dating of business cycle peaks. Time-variation of the transition probabilities is brought about endogenously using the score-driven approach and exogenously using the term spread. In a real-time application using the four components of The Conference Board's Coincident Economic Index for 1959–2020, we find that signaling power for recessions is significantly improved. We are able to date the 2001 and 2008 recession peaks four and two months after the peak date, which is four and ten months before the NBER.

1. Introduction

The business cycle is an important driver of many macroeconomic variables, such as output, employment, etc. Dating the turning points between the phases of this cycle is of utmost importance to policy-makers, firms, and investors. Especially the peaks (marking the transitions from expansions to contractions) of the cycle are of vital interest for downside risk management (Adrian, Boyarchenko, & Giannone, 2019; Caldara, Cascaldi-Garcia, Cuba-Borda, & Loria, 2020). The dynamic factor Markov-switching (DFMS) model proposed by Diebold and Rudebusch (1996) has proven to be a powerful framework to measure the cycle. This model extracts a latent business cycle factor from multiple coincident variables and allows the dynamics of the factor to be regime-dependent using a hidden Markov process. Chauvet and Piger (2008) find that the DFMS model is able to call the troughs (marking the transitions from contractions to expansions) of the cycle faster in real-time than the National Bureau of Economic Research (NBER). However, they find no such improvements in timeliness for the peaks.

This paper addresses the latter issue and extends the DFMS model to accelerate peak dating. With this purpose in mind, we allow for a time-varying transition probability (TVTP) to switch from an expansion to a contraction phase. This probability is a key ingredient for peak dating, and assuming it to be time-invariant may be overly restrictive. For example, one expects the probability of entering a new recession to vary depending on economic fundamentals, e.g., Diebold, Lee, and Weinbach (1994). To bring about time variation, we propose an autoregressive structure driven by the log-likelihood score and, additionally, by leading indicators (LIs). The resulting Generalized Autoregressive Score with eXogenous variables (GASX) model thus combines the ideas of endogenous and exogenous drivers of the transition probabilities from Durland and McCurdy (1994) and Filardo (1994), respectively.

☆ We gratefully acknowledge helpful comments and suggestions from Monica Billio, Francis X. Diebold, Laurent Ferrara, Rutger-Jan Lange, Michel van der Wel and the participants of the CFE-CM Statistics 2019 and SNDE 2020 conferences. Any remaining errors are our own.

* Correspondence to: Burgemeester Oudlaan 50, 3062 PA, Rotterdam, Netherlands.

E-mail addresses: vanos@ese.eur.nl (B. van Os), djvandijk@ese.eur.nl (D. van Dijk).

https://doi.org/10.1016/j.ijforecast.2023.03.005

Please cite this article as: B. van Os and D. van Dijk, Accelerating peak dating in a dynamic factor Markov-switching model. International Journal of Forecasting (2023), https://doi.org/10.1016/j.ijforecast.2023.03.005.

We apply the GASX approach to date US business cycle peaks, based upon the four components of The Conference Board's (TCB) Coincident Economic Index (CEI) for 1959–2020. We consider both an ex-post full-sample analysis using revised data and an ex-ante real-time analysis using data vintages available from December 1976 until March 2020. For the exogenous input, we consider an indicator for a negative term spread, which is generally considered one of the most prominent LIs, see, e.g., Estrella and Mishkin (1998). We find in both ex-post and ex-ante analyses that the GASX model significantly improves upon the DFMS model with static transition probabilities in terms of signaling recessions. Additionally, by converting real-time smoothed state probabilities to turning points, the GASX specification is able to match or precede the peak announcements made by the NBER without any false signals. Most notably, our proposed model is able to date the peaks of the 2001 and 2008 recessions four and two months after the peak date. This is four and ten months before their NBER announcements and a three- and five-month gain over the base DFMS model. In line with the reputation as a powerful LI, most improvements stem from using the term spread. Notably, GAS dynamics are found to amplify correct peak signals further. Therefore, combining both drivers in the GASX model is particularly attractive.

Our paper contributes to the vast literature on US business cycle measurement and dating turning points (e.g., Boldin, 1994; Berge & Jordà, 2011; Hamilton, 2011; Stock & Watson, 2014; Doz, Ferrara, & Pionnier, 2020). In particular, we build upon the DFMS framework of Diebold and Rudebusch (1996), which combines the dynamic factor structure of Stock and Watson (1989, 1993, 2005, 2010) with the Markov-switching (MS) approach of Hamilton (1989). This setup captures the co-movement in multiple coincident series and the regime dependence that characterize the business cycle (Burns & Mitchell, 1946). The factor structure allows for a larger information input relative to univariate MS models, which are already able to effectively date turning points (Layton, 1996; Layton & Katsuura, 2001; Chauvet & Piger, 2003). Furthermore, Chauvet and Piger (2008) find that the DFMS model fares well against the non-parametric dating method of Harding and Pagan (2003), which aggregates the approach of Bry and Boschan (1971) for multiple series. Both methods provide faster real-time dating of the troughs compared to the NBER. Similar results have been found for other countries (Watanabe et al., 2003; Aastveit, Jore, & Ravazzolo, 2016; Carstensen, Heinrich, Reif, & Wolters, 2020).

Our paper closely relates to the literature on macroeconomic MS models with TVTPs. One way to drive these TVTPs is to use exogenous information as LIs. In particular, Filardo (1994) finds that TVTPs driven by LIs can aid in the identification of turning points for output growth. In the context of the DFMS model, Huang and Startz (2020) allow the transition probabilities to depend on the stock market volatility, while Chauvet and Senyuz (2016) add a set of yield curve variables. Both find that turmoil in financial markets often precedes economic recessions. Our proposed GASX model incorporates this insight by allowing the TVTPs to depend on the term spread, which has historically been among the best LIs (e.g., Estrella & Mishkin, 1998; Rudebusch & Williams, 2009; Ng & Wright, 2013; Liu & Mönch, 2016).

Alternatively, one may drive the TVTPs using only endogenous information. This can be done using duration dependence (Durland & McCurdy, 1994; Kim & Nelson, 1998), directly specifying the TVTPs as functions of dependent variables (Diebold et al., 1994; Caldara et al., 2020), or using a score-driven approach (Bazzi, Blasques, Koopman, & Lucas, 2017). The latter updates the TVTPs in the direction of the log-likelihood score as suggested by Creal, Koopman, and Lucas (2013) and Harvey (2013). This Generalized Autoregressive Score-driven (GAS) approach has favorable properties (Blasques, Koopman, & Lucas, 2015), allows for standard likelihood estimation, and is known to produce accurate filters in a variety of settings, see, e.g., Koopman, Lucas, and Scharth (2016). Our GASX approach presents a multivariate version of the GAS setup of Bazzi et al. (2017) and combines it with exogenous LI information in a single framework.

The outline of the paper is as follows. Section 2 presents how the DFMS framework may be enhanced by adding time-varying dynamics to the transition probabilities. Section 3 examines the results of the empirical application both ex-post, with currently available revised data, and in real-time. Section 4 concludes.

2. Methodology

2.1. Model specification

For clarity of exposition, we present the DFMS model for N coincident economic variables with two Markov states, a single factor, and first-order autoregressive (AR(1)) dynamics. Extensions to more regimes, multiple common factors, and higher lag orders are relatively straightforward but tedious. Let $y_{i,t}$ denote the observation of variable $i = 1, 2, \ldots, N$ at time $t = 1, 2, \ldots, T$. We assume that the $y_{i,t}$ are driven by a common latent factor ψ_t with constant factor loadings λ_i and idiosyncratic components $v_{i,t}$, such that the observation equation of the state space representation is given by

$$\mathbf{y}_t = \mathbf{Z}\boldsymbol{\zeta}_t, \tag{1}$$

where $\mathbf{y}_t$ collects the $y_{i,t}$ in a $(N \times 1)$ vector, $\mathbf{Z}$ is a $(N \times N+1)$ matrix of coefficients and $\boldsymbol{\zeta}_t$ is the $(N+1 \times 1)$ state vector, that is,

$$\mathbf{y}_t = \begin{bmatrix} y_{1,t} \\ \vdots \\ y_{N,t} \end{bmatrix}, \ \mathbf{Z} = \begin{bmatrix} \lambda_1 & 1 & 0 & \ldots & 0 \\ \lambda_2 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_N & 0 & 0 & \ldots & 1 \end{bmatrix} \text{ and } \boldsymbol{\zeta}_t = \begin{bmatrix} \psi_t \\ v_{1,t} \\ \vdots \\ v_{N,t} \end{bmatrix}. \tag{2}$$

We assume that the latent factor ψ_t follows a stationary AR(1) process with autoregressive parameter ϕ and intercept α_{S_t}, which depends on the state $S_t \in \{0, 1\}$ of

a hidden inhomogeneous Markov process with dynamic transition probabilities $p_t^{ij} := \Pr(S_t = j|S_{t-1} = i)$ with $i, j \in \{0, 1\}$. For identification purposes, we impose $\alpha_{S_t=0} > \alpha_{S_t=1}$, such that regime 0 (1) reflects an expansion (contraction) period. The variance of the factor innovations is denoted by σ_η^2. We assume stationary zero mean AR(1) dynamics for the idiosyncratic components $v_{i,t}$ with autoregressive parameters θ_i and error variances σ_i^2. The transition equation of the state vector $\boldsymbol{\zeta}_t$ is then given by

$$\boldsymbol{\zeta}_t = \mathbf{d}_{S_t} + \mathbf{V}\boldsymbol{\zeta}_{t-1} + \mathbf{Q}^{\frac{1}{2}}\boldsymbol{\omega}_t, \tag{3}$$

where the system matrices $\mathbf{d}_{S_t}$, $\mathbf{V}$ and $\mathbf{Q}$ are defined as

$$\mathbf{d}_{S_t} = \begin{bmatrix} \alpha_{S_t} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ \mathbf{V} = \begin{bmatrix} \phi & 0 & \dots & 0 \\ 0 & \theta_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta_N \end{bmatrix} \text{ and } \mathbf{Q} = \begin{bmatrix} \sigma_\eta^2 & 0 & \dots & 0 \\ 0 & \sigma_1^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_N^2 \end{bmatrix}, \tag{4}$$

and $\boldsymbol{\omega}_t$ denotes an $(N + 1 \times 1)$ i.i.d. innovation vector, which we assume follows a multivariate standard normal distribution.

2.2. Score-driven time-varying transition probabilities

The transition probabilities of the latent Markov process S_t play a vital role in the timely identification of business cycle turning points. With the aim of accelerating peak dating in mind, we consider dynamic transition probabilities starting from an expansion period (p_t^{00} and $p_t^{01} = 1 - p_t^{00}$) but keep the probabilities starting from a contraction fixed ($p_t^{11} = p^{11}$ and $p_t^{10} = p^{10} = 1 - p^{11}$). For expositional purposes, we formulate our model in terms of p_t^{01}, which reflects the probability of switching from an expansion to a contraction phase, such that we may informally label it as the (conditional) peak probability.

We build upon the general framework of Creal et al. (2013) and use the endogenous information available in the log-likelihood score to drive p_t^{01}. In addition, we also allow for exogenous information as suggested by Diebold et al. (1994) and Filardo (1994). We adopt the logistic link function to ensure the transition probability remains in the unit interval. Specifically, we model the conditional peak probability p_t^{01} as

$$p_t^{01} = \frac{\exp(f_t)}{1 + \exp(f_t)}, \tag{5}$$

$$f_{t+1} = w + as_t + bf_t + cx_t, \tag{6}$$

$$s_t = g(H_t\nabla_t^f), \quad \nabla_t^f = \frac{\partial}{\partial f_t}\log p(y_t|\mathcal{I}_{t-1}), \tag{7}$$

where f_t reflects the log odds ratio of the transition probability p_t^{01} (i.e. $f_t = \log(p_t^{01}/(1 - p_t^{01}))$), x_t is an exogenous variable known at time t, and w, $a > 0$, $b \in (-1, 1)$ and c are static parameters. Furthermore, s_t denotes the endogenous innovation term composed of three components. First and foremost, ∇_t^f denotes the score with respect to f_t and is obtained by taking the derivative of the (approximate) predictive log density at time t, denoted by $\log p(\mathbf{y}_t|\mathcal{I}_{t-1})$. Here $\mathcal{I}_{t-1}$ denotes the information set containing all information available at time $t - 1$. Second, H_t is a positive scaling factor known at time t. Third, $g(\cdot)$ denotes a strictly monotonic transformation with $g(0) = 0$. We note that the true p_t^{01} need not follow the evolution as outlined in (5)–(7). It is, therefore, perhaps more appropriate to view our setup as an intuitive filter that tracks the true time-varying parameter using a steepest ascent-type search. Our setup can be seen as a generalization of Bazzi et al. (2017), which shows that score-driven TVTPs yield an effective filter for various data-generating processes in a univariate setting.

Popular choices for H_t in the GAS literature are no scaling ($H_t = 1$) or powers of the Fischer matrix to account for the curvature of the likelihood. We consider a scaling that allows for more variation of p_t^{01}. Specifically, the score ∇_t^f contains the term $p_t^{01}(1 - p_t^{01})$ because of the logistic link due to the chain-rule. This term is often close to 0, dampening and delaying movement in p_t^{01}. We therefore suggest to use $H_t = 1/(p_t^{01}(1 - p_t^{01}))$ to remove this effect. It is straightforward to show that this part would disappear when scaling with the square root Fischer information. The Fischer information, however, presents a large computational burden here. For $H_t = 1/(p_t^{01}(1 - p_t^{01}))$ and using the likelihood approximation of Kim (1994) our scaled score $H_t\nabla_t^f$ is given by

$$H_t\nabla_t^f = \Pr(S_{t-1} = 0|\mathcal{I}_{t-1})\frac{\phi_t^{01}(\mathbf{y}_t|\mathcal{I}_{t-1}) - \phi_t^{00}(\mathbf{y}_t|\mathcal{I}_{t-1})}{p(\mathbf{y}_t|\mathcal{I}_{t-1})}, \tag{8}$$

where $\phi_t^{ij}(\mathbf{y}_t|\mathcal{I}_{t-1})$ denotes the multivariate normal density evaluated in $\mathbf{y}_t$ conditional on all information available at time $t - 1$ and the states being i and j at time $t - 1$ and t, respectively (see Appendix A for details). The scaled score $H_t\nabla_t^f$ thus considers the difference in likelihood between currently being in a contraction or an expansion phase when the previous period is assumed to be an expansion, weighted by the likelihood. This relative difference is then multiplied with the state probability $\Pr(S_{t-1} = 0|\mathcal{I}_{t-1})$, such that we update p_t^{01} more (less) during an expansion (contraction), precisely when the data tells us much (little) about the expansion dynamics.

The addition of the transformation $g(\cdot)$ in (7) can be seen as a score-equivalent special case of the more general quasi score-driven (QSD) framework of Blasques, Francq, and Laurent (2023). The QSD framework outlines a class of models that nests the GAS approach and also allows for different target functions in place of the local log-likelihood to construct s_t. The corresponding target function is obtained by integrating s_t with respect to f_t. The score-equivalence property refers to the fact that by construction $g(\cdot)$ maintains the update direction (i.e. $\text{sgn}(s_t) = \text{sgn}(\nabla_t^f)$). As a result, we preserve the local optimality properties as laid out in Blasques et al. (2015).

We argue that a transformation $g(\cdot)$ that decreases large (absolute) scores, essentially shrinking large updates, may be especially useful for forecasting. In our case,

in particular, we find empirically that the score in the DFMS model can produce a substantial number of outliers. This happens when the likelihood in the denominator in (8) becomes small. As a result, the learning rate parameter a in (6) is underestimated, limiting the ability to drive the transition probability meaningfully. In a standard regression framework with a strictly positive regressor, a straightforward remedy for this issue would be to consider the logarithm instead. However, as the score can be both positive and negative, we propose the following intuitive transformation:

$$g(x) = \text{sgn}(x)\log(1 + |x|), \tag{9}$$

which is a monotonic antisymmetric function about the origin, coinciding with the zero expectation of the score, that is close to the identity map for 'very small' $|x|$ and close to the logarithm for 'large' $|x|$. This transformation, therefore, maintains the direction of the update while shrinking the updates for large absolute scores. This, in turn, prevents over-updating of our transition probability when a very large score, possibly due to an outlier, occurs. Integrating the transformed score reveals we are now essentially applying a gradient update based on a flattened version of the log-likelihood, similar in spirit to the robust-GAS models proposed by Blasques et al. (2023), who consider, for example, Huber-type target functions. We argue that our choice of $g(\cdot)$ is perhaps the most obvious for our purposes, being a natural extension of the logarithm capable of handling both negative numbers and 0. For robustness, we include results in Appendix C.2 using the inverse hyperbolic sine ($g(x) = \log(x + \sqrt{x^2 + 1})$), a popular transformation in regression analysis used to curb extreme values. We find nearly identical results using the transformation in (9). In addition, we provide results when omitting the transformation ($g(x) = x$)) or the scaling ($H_t = 1$) in Appendix C.3.

2.3. Estimation

As the DFMS model contains both latent regimes and a latent factor, estimation uses the Hamilton and Kalman filters. Furthermore, parameter estimation requires either using an approximation of the likelihood or Bayesian methods, see Kim and Nelson (1999). This is because the calculation of the exact likelihood quickly becomes computationally infeasible due to the problem of path dependence, meaning that the value of the factor at time t depends on all previous Markov states. We follow the approach of Chauvet (1998), which uses the filter proposed by Kim (1994), and approximate the likelihood. Specifically, this method proposes a collapsing step to avoid the need to track an ever-increasing number of past states, such that only a modest history of states needs to be considered. This history length must be at least one longer than the highest lag order in the model to maintain sufficient accuracy; see Kim (1994) for further details. We maximize the associated approximate log-likelihood obtained using the prediction-error decomposition to obtain the parameter estimates. Because of the observation-driven nature of the dynamics of p_t^{01}, the likelihood procedure is unaltered relative to the model with static transition probabilities. The complete prediction-update recursion and further details regarding estimation are provided in Appendix A.

3. Empirical application

3.1. Data

We consider the four components of TCB's CEI for the US economy: employees on nonfarm payrolls (EMP), industrial production (IP), manufacturing and trade sales (MAN), and personal incomeless transfer payments (INC). We analyze these four indicators' monthly logarithmic growth rates from January 1959 until February 2020. Vintages for the real-time exercise are obtained from TCB and supplemented with data from Jeremy Piger.[1] Given its specific and unusual features, the COVID-19 period from March 2020 until December 2020 is investigated separately in Appendix E.

Furthermore, we use the interest rate term spread (TS) as an exogenous driver of the time-varying peak probability. Specifically, we use the TS included in TCB's Leading Economic Index (LEI), the 10-year US Treasury rate minus the US Federal Funds rate. We collect the TS from the FRED database. Because of large differences in level over time, we transform the TS to a dummy variable. Specifically, we set our dummy equal to 1 if the TS is negative and 0 otherwise. This balances the contrast between elevated interest rates during the 1980s and the suppressed term premiums after 2008 due to quantitative easing. Empirically, we find that recessions are almost always preceded by periods of an inverted yield curve,[2] i.e., a negative TS. The TS dummy, therefore, presents a simple and intuitive predictor for future peaks.

Results of several robustness checks, whereby the TS is incorporated differently, can be found in Appendix C.4. These figures additionally contain findings in case the LEI is used as the exogenous variable, which also appears to be an effective choice. We prefer the TS over the LEI because the former tends to signal peaks earlier and does not contain revisions. Considering more or a weighted combination of LIs to drive the peak probability is left for future research.

3.2. Model specification details

We use the two-regime single-factor DFMS model with AR(1) specifications for the common factor and the idiosyncratic components, as described in Section 2. Although more involved multi-state models can provide a better in-sample fit (e.g., Sichel, 1994), a two-regime setup appears more appropriate for our goal of peak dating. The magnitudes of the eigenvalues of the correlation matrix of the four coincident indicators, in turn, motivate the choice of a single common factor. AR(1) dynamics are selected mainly not to overcomplicate the already reasonably involved estimation. Our setup is, therefore, similar to the one of Chauvet (1998) for monthly data. Further extensions, including structural breaks to accommodate the Great Moderation (McConnell & Perez-Quiros, 2000) and time-varying mean growth rates (Eo & Kim,

[1] Obtained from https://pages.uoregon.edu/jpiger/research/published-papers/raw-real-time-data.zip, see Appendix B for details.

[2] A graphical illustration is provided in Appendix B.4.

2016; Doz et al., 2020; Eo & Morley, 2022) are left for future research.

The addition of a time-varying peak probability p_t^{01} substantially enhances the already high flexibility of the DFMS model. In a Bayesian context, Filardo and Gordon (1998) impose strong priors on the TVTPs in their MS model for output to prevent the underestimation of recession durations. This underestimation may lead to 'recession rallies,' meaning that specific months with only moderately negative (or even positive) growth rates may be classified in the expansion regime during a recession period. We propose to similarly impose structure by calibrating the static transition probabilities starting from a recession regime (p^{11} and $p^{10} = 1 - p^{11}$). For this, we use a simple count-based estimator using completed NBER recessions. We obtain p^{11} by dividing the total number of recession months minus the number of recessions by the total number of recession months. This approach is readily applicable in real-time[3] and effectively avoids the issue of 'recession rallies' discussed above. For completeness, results without targeting are provided in Appendix C.3.

We consider four (nested) DFMS model specifications for both the ex-post and the real-time analysis. The first is the base model as given in (1) and (3) with a constant peak probability p^{01}. The remaining three consider a time-varying peak probability p_t^{01} using only endogenous information (GAS), only exogenous drivers (Exo), or both (GASX). Specifically, we use (5)-(9) with $c = 0$ for GAS and $a = 0$ for Exo, where x_t is the TS dummy. To identify the common factor, we fix the factor loading of employment $\lambda_{EMP} = 1$.

3.3. Full-sample results

We estimate the DFMS model and our extensions for January 1959 until February 2020 using the data vintage released in March 2020. Hence, this includes all revisions in the coincident variables known at the final date. In Table 1, only key parameter estimates for the four considered model specifications are shown for brevity. The remaining parameter estimates are in Appendix C.1. Furthermore, Table 1 provides an overview of the signaling performance of the filtered state probabilities $\Pr(S_t = 1|\mathcal{I}_t)$ for the NBER recessions. This includes the Area-Under-the-Receiver-Operating-Curve (AUROC), a common measure for evaluating binary classification ability (Berge & Jordà, 2011). Perfect and uninformative classifiers have AUROC values of 1 and 0.5, respectively. In addition, we consider the average contraction state probability during NBER recessions, non-recession periods, and the first month of the recessions, denoted by π^r, π^e and π^p, respectively.

In Table 1, we observe from the log-likelihood and the Akaike information criterion (AIC) that the extended specifications improve upon the base model to different degrees. This includes notable improvements from using the TS as exogenous input and somewhat more modest improvements due to the endogenous GAS dynamics. GASX is the preferred model according to the AIC. In terms of parameter estimates, we find that the parameter of the negative TS indicator c is positive and significant for both the Exo and GASX specifications. The positive sign is in accordance with economic intuition, which suggests a higher probability of a change to a recession state in the face of an inverted yield curve. In addition, the estimates of the autoregressive parameter b suggest that p_t^{01} is highly persistent.

Table 1
Key parameter estimates and NBER recession signaling performance of the DFMS model and extensions.

	Base	GAS	Exo	GASX
w		−0.559	−0.463	−0.553
		(0.218)	(0.167)	(0.183)
a		1.207		0.979
		(0.399)		(0.461)
b		0.839	0.936	0.910
		(0.050)	(0.023)	(0.025)
c			0.533	0.685
			(0.171)	(0.221)
p^{01}	0.017			
	(0.006)			
LogL	−1913.3	−1907.3	−1900.2	−1897.9
k	16	18	18	19
AIC	3858.5	3850.7	3836.5	3833.9
AUROC	0.941	0.950	0.979***	0.978***
π^r	0.647	0.685	0.779	0.802
π^e	0.066	0.063	0.063	0.065
π^r/π^e	9.731	10.832	12.337	12.419
π^p	0.267	0.290	0.528	0.594

Note: This table presents the key parameter estimates for the base model and GAS, Exo, and GASX specifications for EMP, IP, MAN, and INC from January 1959-February 2020. Standard errors are displayed in parentheses, and k denotes the number of parameters. The signaling power of the filtered states for the NBER recessions is evaluated using the AUROC, $\pi^{r(e)}$ represent the average filtered contraction state probability during NBER recession (expansion) periods and π^p the average probability in the first recession month. Finally, a significant difference in the AUROC compared to the base model is indicated with a *, **, and *** for a p-value below 0.10, 0.05, and 0.01, respectively.

While the improvements in log-likelihood due to the GAS dynamics seem modest, LR tests indicate they are significant, irrespective of including the exogenous variable. Specifically, the LR tests for the GAS versus the base specification and the GASX versus the Exo specification reject the null hypothesis at the 5 percent level ($LR = 11.87$, p-value $= 0.003$ and $LR = 4.6179$, p-value $= 0.032$ respectively). We note that in the absence of a formal proof of the asymptotic distribution of the LR test, some care must be taken to interpret these p-values. However, considering the AIC or the individual significance of the relevant parameters corroborates these findings. That is, we find the GAS parameter a to be significant in both GAS and GASX specifications with lower AICs compared to the base and Exo models.

To formally compare the AUROC between the different models, we follow Aastveit, Anundsen, and Herstad (2019) and use a (two-sided) Wald-type test as suggested by DeLong, DeLong, and Clarke-Pearson (1988). Table 1 shows that all three extensions improve the signaling

[3] For the real-time analysis, we update p^{11} the month after an NBER trough announcement, see Appendix B.5 for its evolution. The ex-post analysis uses all recessions in the sample.

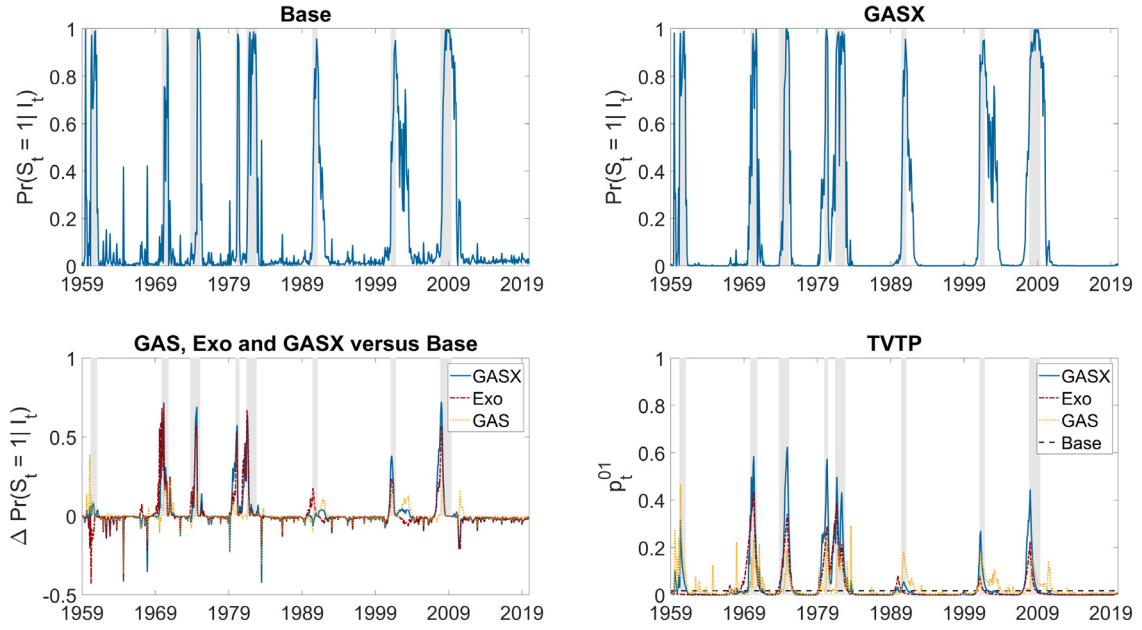

Fig. 1. Filtered state and peak probabilities for the DFMS model and extensions.
Note: The top left (right) plot depicts the filtered state probabilities $\Pr(S_t = 1|\mathcal{I}_t)$ for the base (GASX) model. The bottom left plot contains the filtered state probabilities of the extensions minus those of the base model, denoted by $\Delta\Pr(S_t = 1|\mathcal{I}_t)$. The bottom right plot contains the estimated path of the peak probability p_t^{01}. Finally, the shaded areas reflect the recession periods determined by the NBER.

ability relative to the base model. The largest and most significant improvements stem from the addition of the TS information. For example, this increases the overall AUROC from 0.941 to 0.979 for the Exo specification. While the overall AUROC does not improve when adding GAS dynamics, they benefit peak dating by increasing π^p. In particular, π^p increases from 0.528 to 0.594 in favor of the GASX model relative to the Exo specification. Similarly, the GASX model achieves the highest value of the ratio between π^r and π^e, which can be seen as an informal measure of signal clarity.

Fig. 1 displays the filtered state probabilities $\Pr(S_t = 1|\mathcal{I}_t)$ of the base model and the GASX specification. To disentangle the performance gains of the GASX model, we additionally present the differences of the filtered probabilities of all three extended models (i.e., GAS, Exo, and GASX) with the base model. Positive differences indicate higher recession probabilities in the more extensive model specifications. Finally, Fig. 1 also contains the estimated paths of the peak probability p_t^{01}. A model specification is considered superior if it has a heightened p_t^{01} just before or at the start of a recession.

Fig. 1 shows that the base and the GASX model successfully identify the business cycle regimes. The filtered state probabilities generally remain close to 0 during expansion periods and rapidly increase to levels close to 1 during recessions. Between the two, we observe a clear advantage of the GASX model in terms of recession-signaling ability. In particular, the extended models dramatically increase the contraction probabilities at the start of recessions relative to the base model. In addition, the extended models reduce false signals, such as those in 1964, 1968, and 1983. Comparing the three extended specifications, we observe the largest benefits from including the TS and a secondary but meaningful role for GAS dynamics. In line with the results of Table 1, the combined GASX model offers the best signaling performance.

From the evolution of the dynamic peak probability p_t^{01}, we can more clearly see the differences between the extended models. Specifically, the variation in p_t^{01} for the GAS model appears sensible but is mostly coincident. Therefore, peaks are not signaled in advance; the switches are more extreme. This is unsurprising because the GAS specification still only uses coincident information. We observe large movements before the recessions start for the two other specifications, which include the TS. This timeliness reflects the nature of the TS as a leading indicator. For example, if we currently have that $p_t^{01} = 0.01$, then for the Exo specification, three periods of an inverted yield curve will nearly triple the peak probability to $p_{t+3}^{01} = 0.0274$. Adding GAS dynamics on top further strengthens the movement of p_t^{01} and adds a small increase before the 1960 recession. For this recession, no yield curve signal is contained in the data. We conclude that a time-varying peak probability p_t^{01}, driven by a combination of exogenous and endogenous sources, can help improve the (ex-post) dating of recessions.

3.4. Real-time analysis

We perform a real-time analysis to make a fair assessment of the applicability of the GASX model in practice. Specifically, we recursively estimate the model parameters for each month from December 1976 until March 2020 using the most recent data vintage available. We

Table 2
Real-time signaling performance of the DFMS model and extensions.

Filtered	**Base**	**GAS**	**Exo**	**GASX**	**AIN**
AUROC	0.940	0.944	0.967**	0.970**	0.973*
$\pi^r_{t-1\|t-1}$	0.751	0.783	0.828	0.845	0.609
$\pi^e_{t-1\|t-1}$	0.123	0.119	0.106	0.110	0.121
$\pi^r_{t-1\|t-1}/\pi^e_{t-1\|t-1}$	6.083	6.601	7.846	7.697	5.034
π^p	0.185	0.215	0.387	0.422	0.409
Predicted	**Base**	**GAS**	**Exo**	**GASX**	
AUROC	0.894	0.908**	0.948***	0.955***	
$\pi^r_{t\|t-1}$	0.620	0.669	0.710	0.740	
$\pi^e_{t\|t-1}$	0.140	0.133	0.111	0.116	
$\pi^r_{t\|t-1}/\pi^e_{t\|t-1}$	4.440	5.014	6.423	6.389	
π^p	0.116	0.136	0.241	0.308	

Note: the signaling power of the probabilities for the NBER recessions is evaluated using the AUROC. Furthermore, we have that $\pi^{r(e)}_{i|j}$ represents the average contraction state probability at time i during NBER recession (expansion) periods with observations up to and including time j. This corresponds to the real-time estimation at time $j+1$. In addition, π^p denotes the average state probability during the first month of the recessions within the evaluation sample. The evaluation sample ranges from November 1976 until February 2020. Finally, a significant difference in the AUROC compared to the base model is indicated with a *, **, and *** for a p-value below 0.10, 0.05, and 0.01, respectively.

consider an expanding window approach where the observations started in January 1959. Following Camacho, Perez-Quiros, and Poncela (2018) and in line with TCB's release schedule for their LEI, we estimate the model parameters in the third week of each month. Specifically, this entails that in month t, we have EMP and IP available up to and including month $t-1$. MAN and INC's data publication scheme implies that we have observations only up to and including month $t-3$ and $t-2$, respectively. Parameter estimation in month t is thus based on all observations up to and including month $t-1$, whereby the final two observations for MAN and the final observation for INC are considered missing. Here we differ from Chauvet and Piger (2008), who restrict the sample at each point in time to the series for which the least amount of information is available.

Table 2 presents an overview of the recession signaling performance of the real-time filtered and predicted state probabilities. To assess the economic value of our predictions, we also include the Anxious Index Nowcast (AIN) of Scavette (2014). The AIN is based on the Survey of Professional Forecasters (SPF) and corresponds with the probability of a current contraction in real GDP. To obtain monthly values of the AIN, we set all months within a quarter equal to the value of the AIN in that quarter.

We observe in Table 2 that all extensions that allow for a time-varying peak probability trump the base model in all considered metrics. This includes higher contraction state probabilities during NBER recession periods (π^r) and lower such probabilities during NBER expansion phases (π^e). In addition, the AUROC of the filtered and predicted contraction state probabilities is significantly improved for all but the plain GAS model for the filtered states. Furthermore, the AUROC of the GASX model is on par with that of the AIN (p-value $= 0.770$ for a test of equality). Interestingly, the gain of the AIN over the base model is smaller in a statistical sense, being only significant at the 10 percent level. This is because the variance of the AUROC of the AIN is relatively large. This is in line with the other metrics of Table 2, where the values of $\pi^r_{t-1|t-1}$ indicate that the AIN produces a weaker signal during recessions relative to the GASX model.

Comparing the three extended DFMS models, we find again that the addition of the exogenous information yields the largest improvements. Interestingly, the GASX model improves significantly upon the Exo model in terms of AUROC for the predicted state probabilities (p-value $= 0.032$). Endogenous dynamics are thus significant regardless of the inclusion of the spread. This difference does not reach statistical significance for the filtered state probabilities (p-value $= 0.150$). Even though the absolute differences between the Exo and GASX models are not that large, the estimated standard errors of these differences are very small. This, in turn, leads to large test statistics. In the interest of dating peaks, we find that the average contraction state probability for the first 'official' (NBER) recession month (π^p) is the highest for the GASX model at 0.422 (0.308) for the filtered (predicted) state probabilities. This is more than double the value of the base model at 0.185 (0.116).

Fig. 2 depicts the real-time evolution of the filtered state probabilities for the base model and the GASX specification. This can be understood to reflect the diagonal of the full history of state probabilities found in Appendix D.1. For comparison purposes, we also display the AIN. Additionally, Fig. 2 displays the evolution of the peak probability p^{01}_t.

We observe in Fig. 2 that also, in real-time, the GASX specification matches the NBER recession periods much better than the base model. Compared to the AIN, we find the GASX model to provide the best signal most of the time, while the AIN appears to have better signal troughs. Notably, the trough of the 2001 recession is identified late by the DFMS models as also found by Chauvet and Piger (2008). This may be due to what is known as a jobless recovery, see Groshen and Potter (2003). While the average signaling power of the AIN measured by the AUROC is similar to the GASX model, we still prefer the latter for peak dating. Compared to the base model, the GASX specification can better identify business cycle peaks for four of the five recessions from 1976 to 2019. Only for the 1990 recession, the GASX model initially appears to dismiss the signal relative to the base model slightly. Interestingly, Stock and Watson (1993) found similar results using their leading indicator when incorporating the TS; see Hamilton (2011) for a discussion. Most strikingly, the 2001 and 2008 recessions are identified much earlier by the GASX model than the base model and the AIN.

Having observed the largest benefits for the GASX specification in the two most recent recessions, we further investigate the evolution of the state probabilities around these periods. Fig. 3 depicts the evolution of the real-time smoothed contraction state probabilities for the recessions that followed the peaks in March 2001 and December 2007.

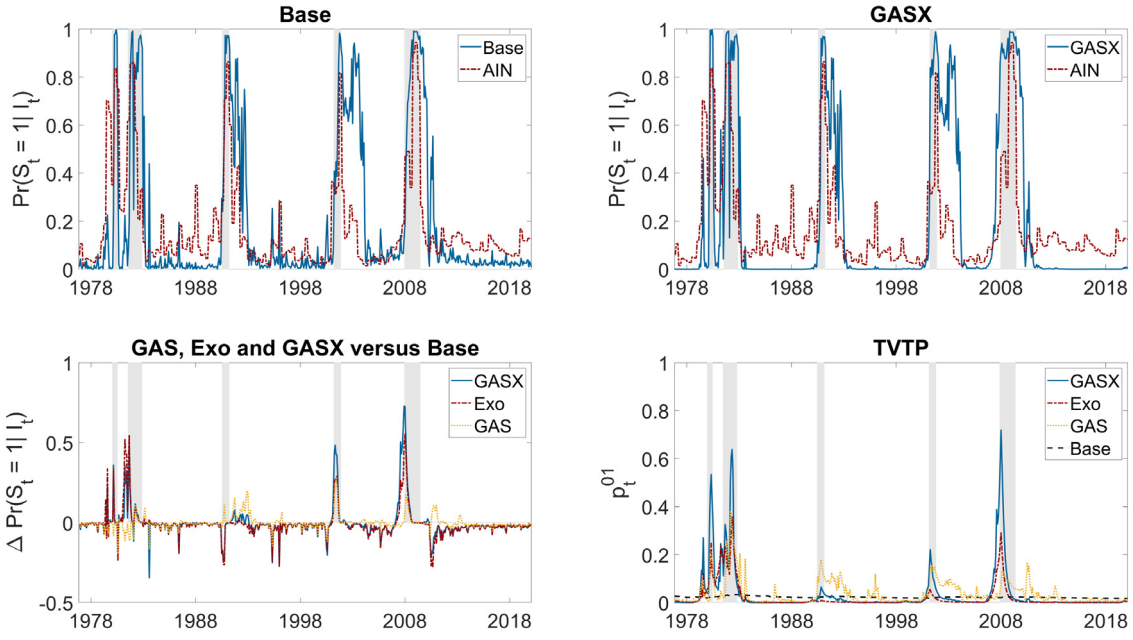

Fig. 2. Real-time results for the base DFMS model and the GASX extension.
Note: The top left (right) figure displays the filtered state probabilities for the base (GASX) model. The bottom left plot displays the difference in filter state probabilities with the base model, while the bottom right figure contains the evolution of the peak probability p_t^{01}. Finally, the shaded areas reflect the recession periods determined by the NBER.

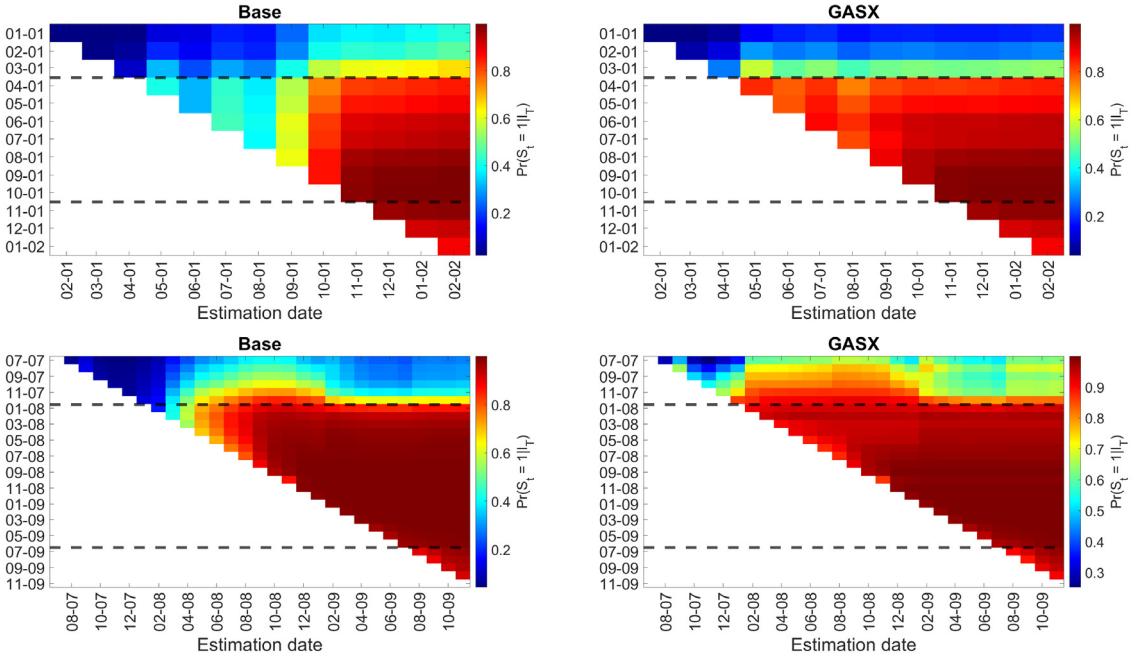

Fig. 3. History of real-time smoothed state probabilities for the base DFMS model and the GASX extension around the 2001 and 2008 recession.
Note: This figure depicts the real-time smoothed state probabilities around the 2001 (top) and 2008 (bottom) recessions for the base (left) and GASX specification (right). The *x*-axis contains the estimation date, and the *y*-axis is the sample date for which a probability is constructed. Finally, the black dotted lines reflect the turning points determined by the NBER.

In Fig. 3, the GASX specification produces much higher contraction state probabilities than the base model during the first months of the 2001 and 2008 recessions. As a result, the GASX specification can accurately provide a strong contraction signal as soon as these recessions start. In contrast, the base model requires several months of additional information to do so. As time passes, the models largely agree on past states. Appendix D.2 contains a plot

similar to Fig. 3 for the filtered state probabilities. Here improvements are even more pronounced. All in all, these findings indicate that the GASX specification may be able to date the 2001 and 2008 peaks several months ahead of the base model, depending on the conversion rule, which we consider next.

To identify peak dates, we use a straightforward two-step conversion rule similar to the one used by Chauvet and Piger (2008). Specifically, we propose the following identification scheme for each estimation date (i.e., for each vintage 'column'). First, to call a recession we require $\Pr(S_t = 1|\mathcal{I}_T) < \tau$ and $\Pr(S_{t+k} = 1|\mathcal{I}_T) \geq \tau$, for $k = 1, 2, 3$ for some threshold $0.5 < \tau < 1$, where T is the last month of the vintage. Second, the peak associated with this recession period is identified by the point in time where immediately preceding this period, the probability crosses 0.5. That is, we find the smallest non-negative integer q such that $\Pr(S_{t-q} = 1|\mathcal{I}_T) < 0.5$ and $\Pr(S_{t-q+1} = 1|\mathcal{I}_T) \geq 0.5$. The peak date is subsequently identified to be month $t - q$ and, as such, refers to the final expansion month. Before another peak can be called, we require that the state probability remains below τ for three consecutive periods. The first such a period after a peak may then be established to be its corresponding trough, although more generally, it may be given its threshold. Here for simplicity, we label the first month t, after a peak has been dated and for which $\Pr(S_t = 1|\mathcal{I}_T) \geq \tau$ and $\Pr(S_{t+k} = 1|\mathcal{I}_T) < \tau$, for $k = 1, 2, 3$ as a trough. The trough, therefore, marks the final recession month. By construction, we now have that peaks and troughs alternate. By considering the recent state probabilities of the initial estimation date, December 1976, we determine that our evaluation window begins in an expansion. We, therefore, begin looking for a peak. In Table 3, the initial peak dates obtained from the abovementioned method for $\tau = 0.65$ and $\tau = 0.8$ are presented and reflect identification in real-time.

In Table 3, we observe that the DFMS models are able to match or precede the peak announcement dates of the NBER, often by a substantial margin. Specifically, the GASX specification is able to date the 2001 and 2008 peaks 4 and 11 (10) months earlier than the NBER for the threshold $\tau = 0.65$ (0.80). This is a gain of three and four (five) months over the base model. The large gain of the GASX model in timeliness relative to the NBER, therefore, appears to be about half due to the GASX structure and half due to the DFMS model itself. We find comparable performance of the base model and the extension for the first three recessions but note that both can generally match or precede the NBER announcements. Of course, the NBER has a different aim with its dating procedure, valuing accuracy above speed. In terms of the peak dates, we find that the DFMS model produces similar dates as the NBER for the first four recessions. The peak date of the 2008 recession, however, is initially identified by the GASX model 4 (6) months before the corresponding NBER peak date for the threshold $\tau = 0.65$ (0.80). This is consistent with Fig. 3, where the smoothed contraction state probabilities are elevated above 0.5 several months before the NBER peak date.

Because the dating procedure is done at each point in time (i.e., for each 'column'), it might be that later estimation dates produce different turning points than initially established. The values in parentheses in Table 3 reflect the peak dates determined at the final estimation date of March 2020. Here we observe that the DFMS specifications make some adjustments as more data becomes available, unlike the NBER, which has not made any revisions since the inception of its dating method. For example, the 2008 recession peak is dated closer to the date established by the NBER at the final estimation date for both specifications. This suggests a trade-off between timeliness and accuracy when comparing our approach with the NBER's.

Table 3
Comparison peak dates from the DFMS specifications with the NBER recessions.

$\tau = 0.65$					
Peak date			*Announcement date*		
Base	**GASX**	**NBER**	**Base**	**GASX**	**NBER**
0 (1)	1 (0)	Jan 1980	0	0	Jun 1980
0 (0)	0 (0)	Jul 1981	−1	−2	Jan 1982
−1 (−2)	−1 (−2)	Jul 1990	−4	−4	Apr 1991
−1 (−2)	−1 (−1)	Mar 2001	−1	−4	Nov 2001
−2 (0)	−4 (−1)	Dec 2007	−7	−11	Dec 2008
$\tau = 0.8$					
Peak date			*Announcement date*		
Base	**GASX**	**NBER**	**Base**	**GASX**	**NBER**
0 (1)	1 (0)	Jan 1980	0	1	Jun 1980
0 (0)	1 (0)	Jul 1981	−1	0	Jan 1982
−1 (−2)	−1 (−2)	Jul 1990	−4	−4	Apr 1991
−1 (−2)	−1 (−1)	Mar 2001	−1	−4	Nov 2001
−5 (0)	−6 (−1)	Dec 2007	−5	−10	Dec 2008

Note: This table contains the monthly differences in obtained initial peak dates of the base model and GASX extension with the NBER database. The numbers in parentheses reflect the dating at the final estimation date of March 2020. Peaks are constructed from the smoothed contraction state probabilities using a threshold of $\tau = 0.65$ (top) and $\tau = 0.8$ (bottom). The NBER turning points and their respective announcement dates are obtained from https://www.nber.org/cycles.html.

With regards to the troughs corresponding to the peaks of Table 3, we confirm the results of Chauvet and Piger (2008). That is, the DFMS model is generally able to call troughs much earlier than the NBER, with the exception of the 2001 recession associated with a jobless recovery. Doz et al. (2020) find that opting for a different measure of employment (specifically, civilian unemployment) may remedy this dating delay. Differences between the base DFMS model and the GASX specification are small. This finding is unsurprising as the transition probability p^{11} plays a much larger role here than p_t^{01} and is set to the same value for both specifications. For this reason and for brevity, the troughs corresponding to the peaks of Table 3 can be found in Appendix D.3.

We conclude that qualitatively our findings in real-time closely match those of our ex-post analysis. This entails that the GASX model provides clear benefits over the base DFMS model in terms of recession-signaling ability. In particular, the GASX model is able to identify peaks faster in real time than the NBER.

4. Concluding remarks

We analyze the possibilities to accelerate the dating of business cycle peaks by extending the DFMS model of Diebold and Rudebusch (1996) to allow for TVTPs. Specifically, we propose using a score-driven approach to guide the transition probabilities based on the techniques found in Bazzi et al. (2017). Additionally, we allow for relevant exogenous economic drivers, as suggested by Filardo (1994). Using the components of TCB's CEI from 1959 until 2020 and the term spread as an exogenous input, we allow the transition probability to switch from an expansion to a contraction phase to be time-varying. The proposed method can significantly accelerate the real-time peak dating for US recessions. The most benefit is obtained from using the term spread, with a more limited but economically meaningful role for score-driven dynamics. Therefore, we recommend a combined approach, such as the GASX model, to predict whether a recession is approaching.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2023.03.005.

This appendix contains additional details regarding estimation (part A) and data (part B). In addition, it contains several robustness checks and supplementary results for the full-sample analysis and real-time analysis (part C and D, respectively). Finally, the effects of the COVID-19 period are investigated (part E).

References

Aastveit, K. A., Anundsen, A. K., & Herstad, E. I. (2019). Residential investment and recession predictability. *International Journal of Forecasting*, *35*(4), 1790–1799.

Aastveit, K. A., Jore, A. S., & Ravazzolo, F. (2016). Identification and real-time forecasting of Norwegian business cycles. *International Journal of Forecasting*, *32*(2), 283–292.

Adrian, T., Boyarchenko, N., & Giannone, D. (2019). Vulnerable growth. *American Economic Review*, *109*(4), 1263–1289.

Bazzi, M., Blasques, F., Koopman, S. J., & Lucas, A. (2017). Time-varying transition probabilities for Markov regime switching models. *Journal of Time Series Analysis*, *38*(3), 458–478.

Berge, T. J., & Jordà, Ò. (2011). Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal: Macroeconomics*, *3*(2), 246–277.

Blasques, F., Francq, C., & Laurent, S. (2023). Quasi score-driven models. *Journal of Econometrics*, *234*(1), 251–275.

Blasques, F., Koopman, S. J., & Lucas, A. (2015). Information-theoretic optimality of observation-driven time series models for continuous responses. *Biometrika*, *102*(2), 325–343.

Boldin, M. D. (1994). Dating turning points in the business cycle. *Journal of Business*, *67*(1), 97–131.

Bry, G., & Boschan, C. (1971). Standard business cycle analysis of economic time series. In *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs* (pp. 64–150). NBER.

Burns, A. F., & Mitchell, W. C. (1946). The basic measures of cyclical behavior. In *Measuring Business Cycles* (pp. 115–202). NBER.

Caldara, D., Cascaldi-Garcia, D., Cuba-Borda, P., & Loria, F. (2020). Understanding growth-at-risk: a Markov-switching approach. Philadelphia Fed Conference on Real-Time Data Analysis, Methods and Applications.

Camacho, M., Perez-Quiros, G., & Poncela, P. (2018). Markov-switching dynamic factor models in real time. *International Journal of Forecasting*, *34*(4), 598–611.

Carstensen, K., Heinrich, M., Reif, M., & Wolters, M. H. (2020). Predicting ordinary and severe recessions with a three-state Markov-switching dynamic factor model: an application to the German business cycle. *International Journal of Forecasting*, *36*(3), 829–850.

Chauvet, M. (1998). An econometric characterization of business cycle dynamics with factor structure and regime switching. *International Economic Review*, *39*(4), 969–996.

Chauvet, M., & Piger, J. (2003). Identifying business cycle turning points in real time. *Federal Reserve Bank of St. Louis Review*, *85*(2), 47.

Chauvet, M., & Piger, J. (2008). A comparison of the real-time performance of business cycle dating methods. *Journal of Business & Economic Statistics*, *26*(1), 42–49.

Chauvet, M., & Senyuz, Z. (2016). A dynamic factor model of the yield curve components as a predictor of the economy. *International Journal of Forecasting*, *32*(2), 324–343.

Creal, D., Koopman, S. J., & Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, *28*(5), 777–795.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, *44*, 837–845.

Diebold, F. X., Lee, J.-H., & Weinbach, G. C. (1994). Regime switching with time-varying transition probabilities. *Business Cycles: Durations, Dynamics, and Forecasting*, *1*, 144–165.

Diebold, F. X., & Rudebusch, G. D. (1996). Measuring business cycles: a modern perspective. *The Review of Economics and Statistics*, *78*(1), 67–77.

Doz, C., Ferrara, L., & Pionnier, P.-A. (2020). Business cycle dynamics after the great recession: an extended Markov-switching dynamic factor model. OECD Statistics Working Papers 2020/01.

Durland, J. M., & McCurdy, T. H. (1994). Duration-dependent transitions in a Markov model of US GNP growth. *Journal of Business & Economic Statistics*, *12*(3), 279–288.

Eo, Y., & Kim, C.-J. (2016). Markov-switching models with evolving regime-specific parameters: Are postwar booms or recessions all alike? *The Review of Economics and Statistics*, *98*(5), 940–949.

Eo, Y., & Morley, J. (2022). Why has the US economy stagnated since the Great Recession? *The Review of Economics and Statistics*, *104*(2), 246–258.

Estrella, A., & Mishkin, F. S. (1998). Predicting US recessions: Financial variables as leading indicators. *The Review of Economics and Statistics*, *80*(1), 45–61.

Filardo, A. J. (1994). Business-cycle phases and their transitional dynamics. *Journal of Business & Economic Statistics*, *12*(3), 299–308.

Filardo, A. J., & Gordon, S. F. (1998). Business cycle durations. *Journal of Econometrics*, *85*(1), 99–123.

Groshen, E. L., & Potter, S. (2003). Has structural change contributed to a jobless recovery? *Current Issues in Economics and Finance*, *9*(8).

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*(2), 357–384.

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

Harding, D., & Pagan, A. (2003). A comparison of two business cycle dating methods. *Journal of Economic Dynamics & Control*, *27*(9), 1681–1690.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press.

Huang, Y.-F., & Startz, R. (2020). Improved recession dating using stock market volatility. *International Journal of Forecasting*, *36*(2), 507–514.

Kim, C.-J. (1994). Dynamic linear models with Markov-switching. *Journal of Econometrics*, *60*(1–2), 1–22.

Kim, C.-J., & Nelson, C. R. (1998). Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. *The Review of Economics and Statistics*, *80*(2), 188–201.

Kim, C.-J., & Nelson, C. R. (1999). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. The MIT Press.

Koopman, S. J., Lucas, A., & Scharth, M. (2016). Predicting time-varying parameters with parameter-driven and observation-driven models. *The Review of Economics and Statistics*, *98*(1), 97–110.

Layton, A. P. (1996). Dating and predicting phase changes in the US business cycle. *International Journal of Forecasting*, *12*(3), 417–428.

Layton, A. P., & Katsuura, M. (2001). Comparison of regime switching, probit and logit models in dating and forecasting US business cycles. *International Journal of Forecasting*, *17*(3), 403–417.

Liu, W., & Mönch, E. (2016). What predicts US recessions? *International Journal of Forecasting*, *32*(4), 1138–1150.

McConnell, M. M., & Perez-Quiros, G. (2000). Output fluctuations in the United States: What has changed since the early 1980's? *American Economic Review*, *90*(5), 1464–1476.

Ng, S., & Wright, J. H. (2013). Facts and challenges from the Great Recession for forecasting and macroeconomic modeling. *Journal of Economic Literature*, *51*(4), 1120–1154.

Rudebusch, G. D., & Williams, J. C. (2009). Forecasting recessions: the puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics*, *27*(4), 492–503.

Scavette, A. (2014). Are we in a recession? The 'anxious index nowcast' knows!. Federal Reserve Bank of Philadelphia *Special Report*.

Sichel, D. E. (1994). Inventories and the three phases of the business cycle. *Journal of Business & Economic Statistics*, *12*(3), 269–277.

Stock, J. H., & Watson, M. W. (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, *4*, 351–394.

Stock, J. H., & Watson, M. W. (1993). A procedure for predicting recessions with leading indicators: econometric issues and recent experience. *Business Cycles, Indicators and Forecasting*, *28*, 95–156.

Stock, J. H., & Watson, M. W. (2005). Implications of dynamic factor models for VAR analysis. NBER Working Paper 11467.

Stock, J. H., & Watson, M. W. (2010). Indicators for dating business cycles: cross-history selection and comparisons. *American Economic Review*, *100*(2), 16–19.

Stock, J. H., & Watson, M. W. (2014). Estimating turning points using large data sets. *Journal of Econometrics*, *178*(2), 368–381.

Watanabe, T., et al. (2003). Measuring business cycle turning points in Japan with a dynamic Markov switching factor model. *Monetary Econ. Stud.*, *21*(1), 35–68.