Accelerating Peak Dating in a Dynamic Factor Markov-Switching Model

Bram van Os
Dick van Dijk

1 Econometric Institute, Erasmus University Rotterdam
Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and Vrije Universiteit Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at https://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Accelerating Peak Dating in a
Dynamic Factor Markov-Switching
Model

BRAM VAN OS* AND DICK VAN DIJK†

Econometric Institute, Erasmus University Rotterdam

August 31, 2020

Abstract The dynamic factor Markov-switching (DFMS) model introduced by Chauvet (1998) has proven to be a powerful framework to measure the business cycle. We extend the DFMS model by allowing for time-varying transition probabilities, with the aim of accelerating the real-time dating of turning points between expansion and recession regimes. Time-variation of the transition probabilities is brought about endogenously using the accelerated 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 the period 1959-2020, we find that signaling power for recessions is significantly improved.

*vanos@ese.eur.nl
†djvandijk@ese.eur.nl
1 Introduction

The business cycle is an important driver of many macroeconomic variables. Dating the turning points between the phases of this cycle, especially the peaks, is of great interest to policy-makers, firms and investors alike. The dynamic factor Markov-switching (DFMS) model proposed by Chauvet (1998) has proven to be a powerful framework to measure the cycle, see Chauvet and Piger (2008) among others. This model extracts a latent business cycle factor by exploiting the cross-sectional information in multiple observed coincident variables. Furthermore, in line with the macroeconomic intuition of expansion and contraction phases in the cycle, the evolution of the factor is allowed to be regime-dependent with a hidden Markov process dictating the regime-switches. Chauvet and Piger (2008) find that the DFMS model compares favorably to the non-parametric dating method of Harding and Pagan (2003). Moreover, they find that the DFMS model is able to call the troughs of the cycle faster in real-time than the National Bureau of Economic Research (NBER), but do not find similar levels of increased timeliness for the peaks.

In this paper we extend the DFMS model of Chauvet (1998) with the aim of accelerating peak dating. With this purpose in mind, we allow the probability to switch from an expansion to a contraction phase to be time-varying. To bring about such time-variation we propose an autoregressive structure driven by endogenous information in the form of the log-likelihood score and additionally allow for exogenous variables. This framework is therefore a multivariate extension of the methods described by Bazzi et al. (2017) for Markov-switching models for univariate time series. Furthermore, we make use of the accelerated score-driven method recently introduced by Blasques et al. (2019) and tailor it to our needs. In this approach not only the magnitude of the score, but also its alignment with its predecessors is considered in the size of the parameter update. Our main contribution is the resulting accelerated Generalized Autoregressive Score with eXogenous variables (aGASX) model. This framework thus combines the ideas of exogenous and endogenous drivers of the transition probabilities from Diebold et al. (1994) and Durland and McCurdy (1994), respectively. In
the context of business cycle turning point dating, the proposed framework provides a simple and more directed alternative to more involved multi-factor and multi-regime approaches.

We explore the empirical usefulness of the aGASX framework in an application involving the four components of The Conference Board’s (TCB) Coincident Economic Index (CEI) for the period 1959-2020. We consider both an ex-post analysis using the full sample with currently available revised data and a real-time exercise using appropriate vintages available from December 1976 until March 2020. Our main empirical findings can be summarized as follows. First, in the ex-post analysis we find that both score-driven endogenous dynamics and exogenous time-variation are significant and improve the signaling ability for recessions. For the exogenous inputs we consider an indicator for a negative term spread, constructed by subtracting the Federal Funds rate from the 10-year Treasury rate, which is generally considered to be one of the most prominent leading indicators, see e.g. Estrella and Mishkin (1998). While most increased dating performance stems from the exogenous information, we find that (accelerated) GAS dynamics are able to meaningfully amplify correct peak signals and reduce false ones. Second, we find that the aGASX specification provides similar benefits over the DFMS model with time-invariant transition probabilities in real-time. This includes a significantly improved signaling power of the real-time contraction state probabilities for the NBER recession periods, with higher such probabilities at the start of four out five of the most recent recessions. Additionally, by converting real-time smoothed state probabilities to turning points, the DFMS-aGASX 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 associated with the 2001 and 2008 recessions four and ten months before their NBER announcements, a gain of three and five month over the DFMS model with time-invariant transition probabilities, respectively.

Our paper is related to several strands of literature. First, it is related to the vast literature on business cycle measurement. Here factor models have played and continue to play a prominent role. Stock and Watson (1989) propose a dynamic factor model and estimate
both a leading and a coincident economic index using a selection of economic indicators. The DFMS model by Chauvet (1998) combines the idea of co-movement in multiple coincident macroeconomic series with the idea of regime-dependence, in line with the notion of expansions and contraction phases in the cycle as found by Hamilton (1989). Chauvet and Piger (2008) find that the DFMS model provides superior timeliness in terms of the real-time dating of turning points when compared to the NBER and the non-parametric method of Harding and Pagan (2003), mainly for the troughs. For an overview and comparison with more alternative dating methods, see Hamilton (2011). The DFMS model has also been successfully applied to macroeconomic data of many other countries, see e.g. Norway by Aastveit et al. (2016), Germany by Carstensen et al. (2020) and Japan by Watanabe et al. (2003) among others, and today still remains a topic of interest. Recently, for example, Camacho et al. (2018) investigate the effects of ragged edges for the DFMS model. The transition probabilities play a primary role in the construction of the regime probabilities. In the context of empirical macroeconomics, Diebold et al. (1994) argue that the transition probabilities of a Markov-switching model need not be constant and propose using economic variables to guide their evolution. Similarly, Filardo et al. (1998) allow transition probabilities to vary over time using a leading economic index in a Bayesian framework. Furthermore, evidence of duration dependence is found by Durland and McCurdy (1994) for GNP growth rates and by Kim and Nelson (1998), using a DFMS model, for coincident indicators.

Second, our paper is related to the literature that has incorporated leading indicators (LIs) in attempts to improve business cycle measurement, see Marcellino (2006) for an overview of the use and construction of LIs. Notably, Paap et al. (2009) find using a Markov-switching vector autoregressive model that the TCB’s Composite Leading Index (LEI) leads the TCB’s CEI by almost a full year for the peaks and a quarter for the troughs. In the context of the DFMS model, Huang and Startz (2018) allow the transition probabilities to depend on the volatility regime of the stock market. In a similar spirit Chauvet and Senyuz (2016) add a set of yield curve variables to the DFMS framework and consider a bi-factor
setup. Here both find that turmoil in the financial markets often precedes recessions of the real economy. While it is incredibly difficult, if not impossible, to pick a single best exogenous predictor for the entire history of the US business cycle, due to e.g. altered drivers of the business cycle, the term spread has historically been a good predictor, see Estrella and Mishkin (1998), Rudebusch and Williams (2009), Ng and Wright (2013) and Liu and Mönch (2016) among many others.

Third, our paper is related to the literature regarding Generalized Autoregressive Score-driven (GAS) models, proposed by Creal et al. (2013). This method updates a time-varying parameter in the direction of the score and thus exploits the information contained in the entire density. At the same time, independently, Harvey (2013) developed Dynamic Conditional Score models using similar methods. Koopman et al. (2016) show in an extensive Monte Carlo study that GAS models offer comparable performance to parameter-driven models, even if the latter matches the data generating process. Due to the observation-driven approach, simple maximum likelihood may be employed for estimation. Moreover, the approach has been found useful in a variety of empirical applications and provides theoretical benefits in terms of local Kullback-Leibler divergence. In the context of Markov-switching models in particular, Bazzi et al. (2017) propose letting the transition probabilities vary over time using the score-driven approach and find improvements for the variance regimes of industrial production growth rates. Recently, Blasques et al. (2019) have suggested reapplying the score-driven approach to the parameter of the score, resulting in the so-called accelerated GAS (aGAS) specification. This extension considers the recent alignment of scores with their predecessors and updates more in the case of higher alignment.

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. Finally, section 4 concludes.
2 Methodology

2.1 Model specification

We present the DFMS model for $N$ coincident economic variables with two Markov states and a single factor. Extensions to more regimes and more common factors are, however, relatively straightforward. Let $y_{i,t}$ denote the observation of variable $i = 1, 2, ..., N$ at time $t = 1, 2, ..., T$. We assume that the $y_{i,t}$ are driven by a common latent factor $F_t$ with factor loadings $\lambda_i$. The latent factor is assumed to follow an autoregressive process of order $p$ (AR($p$)), with autoregressive parameters $\phi_1, \ldots, \phi_p$ such that all roots of the AR-polynomial are outside the unit circle. Here the intercept $\alpha_{S_t}$ is allowed to depend on the hidden Markov state denoted by $S_t \in \{0, 1\}$. Furthermore, $\eta_t$ denotes the shock to the common factor and the idiosyncratic components are then denoted by $v_{i,t}$. The latter are assumed to follow an AR($q$) process with autoregressive parameters $\theta_{i,1}, \ldots, \theta_{i,q}$, such that all roots of the AR-polynomials are outside the unit circle, and innovations $\varepsilon_{i,t}$. The exact model specification is given as follows for each series $i$ at time point $t$:

$$y_{i,t} = \lambda_i F_t + v_{i,t}, \quad (1)$$

$$F_t = \alpha_{S_t} + \sum_{j=1}^{p} \phi_j F_{t-j} + \eta_t, \quad (2)$$

$$v_{i,t} = \sum_{j=1}^{q} \theta_{i,j} v_{i,t-j} + \varepsilon_{i,t}. \quad (3)$$

Here we assume for each time $t$ that $\eta_t$ and $\varepsilon_{i,t}$, $i = 1, 2, ..., N$ follow i.i.d. joint multivariate normal distributions with a diagonal covariance matrix and variances $\sigma^2_{\eta}$ and $\sigma^2_{\varepsilon}$, that is

$$\begin{bmatrix} \eta_t \\ \varepsilon_{1,t} \\ \vdots \\ \varepsilon_{N,t} \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} 0 \\ \sigma^2_{\eta} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{\eta} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\varepsilon} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2_{\varepsilon_N} \end{bmatrix}). \quad (4)$$
The transition probabilities of the latent Markov process play a key role in the timely identification of business cycle turning points. With the aim of speeding up this dating process, our novelty therefore pertains to altering the evolution dynamics of the Markov chain, by allowing its transition probabilities to vary over time. This is achieved by extending the methods found in Bazzi et al. (2017) who consider time-varying transition probabilities in a univariate Markov-switching model, by applying the generalized autoregressive score-driven (GAS) approach from Creal et al. (2013). We extend their work first by considering not a single but multiple time-series by virtue of the factor structure found in the DFMS model. Second, in a similar spirit as Diebold et al. (1994), we allow for relevant economic exogenous variables to guide the transition probabilities. Third, we also explore the novel accelerated GAS approach (aGAS) described in Blasques et al. (2019) and adapt it to our needs. This extension allows for more rapid changes from score information in the time-varying parameter in question if the scores align with their immediate predecessor and smaller changes if this is not the case.

Let $p_{ij}^t$ denote the transition probability $Pr(S_t = j|S_{t-1} = i)$ with $i, j \in \{0, 1\}$. To ensure that the transition probabilities remain in their appropriate interval we consider the following link function

$$
p_{kk}^t = \frac{\exp(f_{k,t})}{1 + \exp(f_{k,t})}, \quad k \in \{0, 1\}.
$$

(5)

Now assume that the time-variation of the vector $f_t = [f_{0,t}, f_{1,t}]^T$ is captured by the following first-order autoregressive process:

$$
f_{t+1} = w + As_t + Bf_t + CX_t.
$$

(6)

We follow the general paradigm of Creal et al. (2013) and consider the score function to construct $s_t$. Note that within the class of observation-driven time-varying parameter models the vector $s_t$ is principle allowed to be any known function of observables. Additionally, let $X_t$ denote some vector of exogenous variables known at time $t$. Here $w$ denotes a vector of...
constants and \( A, B \) and \( C \) are matrices of coefficients. Note also that extensions to a higher number of regimes here are straightforward, but more cumbersome as it requires one to also describe the evolution equations of the off-diagonal elements of the transition matrix.

To facilitate further notation, mainly for the likelihood and score inference in the next section, we assume first-order autoregressive processes for both the factor and the idiosyncratic components, i.e. \( p = q = 1 \) in Equation (2) and (3). In fact for our empirical application this choice of lag-order will suffice. In this case the model specification can be simplified and written in the state space representation below.

\[
y_t = Z \zeta_t
\]

\[
\zeta_t = d_{S_t} + T \zeta_{t-1} + Q^{1/2} \omega_t,
\]

whereby \( y_t \) are the \( y_{i,t} \) collected in a column vector and \( \zeta_t \) is the state vector i.e.

\[
y_t = \begin{bmatrix} y_{1,t} \\ \vdots \\ y_{N,t} \end{bmatrix} \quad \text{and} \quad \zeta_t = \begin{bmatrix} F_t \\ v_{1,t} \\ \vdots \\ v_{N,t} \end{bmatrix}.
\]

The system matrices \( Z, T, d_{S_t}, \) and \( Q \) are defined as

\[
Z = \begin{bmatrix} \lambda_1 & 1 & 0 & \ldots & 0 \\ \lambda_2 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ldots & \vdots \\ \lambda_N & 0 & 0 & \ldots & 1 \end{bmatrix}, \quad T = \begin{bmatrix} \phi & 0 & \ldots & 0 \\ 0 & \theta_1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \theta_N \end{bmatrix},
\]

\[
d_{S_t} = \begin{bmatrix} \alpha_{S_t} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} \sigma_{\eta}^2 & 0 & \ldots & 0 \\ 0 & \sigma_1^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_N^2 \end{bmatrix}.
\]

Here \( \omega_t \) denotes an i.i.d. innovation vector which is assumed to follow a multivariate standard normal distribution of dimension \( N + 1 \). Note also that using the notation above the
specification can in principle be easily generalized to encompass AR(\(p\)) processes, as any AR(\(p\)) can be written as a VAR(1) process in companion form. In that case the factor \(F_t\) would become a vector \(F_t\) consisting of \(F_t\) and \(p-1\) of its lags, whereby the system matrices need be adjusted accordingly.

### 2.2 Likelihood and score

As the DFMS model contains both latent regimes and a latent factor, estimation requires either the use of 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, as the value of the factor depends on all previous states, a problem known as path-dependence. We follow the approach of Kim (1994) and make use of an approximation of the likelihood. Specifically, this method proposes a collapsing step to prevent an increasing amount of states that needs to be tracked. Using this approach only a small history of states needs to be considered, such that the conditional density at each point in time remains a normal mixture with a tractable number of terms. To maintain sufficient accuracy one requires this history length to be at least one higher than the highest lag-order contained in the model, see Kim (1994) for further details. The approximated log-likelihood to be optimized can be obtained using the prediction-error decomposition. For our current setting with \(p = q = 1\) this results in

\[
\log L \approx \sum_{t=1}^{T} \log \left( \sum_{i,j \in \{0,1\}} p_{ij} \Pr(S_{t-1} = i | I_{t-1}) \phi_{ij}(y_t) \right). 
\]  

Here \(\phi_{ij}(y_t)\) denotes the multivariate normal density evaluated in \(y_t\) with mean \(\mu_{ij}\) and covariance \(\Sigma_{ij}\), in turn defined as

\[
\mu_{ij} = Z \zeta_{ij|t-1},
\]

\[
\Sigma_{ij} = Z P_{ij|t-1} Z'.
\]
Where $\zeta_{ij}^{t|t-1}$ denotes the expectation of the state vector $\zeta_t$ conditional on the information set $I_{t-1}$ and the states being $i$ and $j$ at time $t-1$ and $t$ respectively, similarly define $\mathbf{P}_{ij}^{t|t-1}$ to be its covariance matrix.

The gradient or score of the (approximated) predictive log density at time $t$ with respect to the $f_{k,t}$, which make up the transition probabilities, is denoted by $\nabla f_t$ and is given as follows:

$$\nabla f_t = \frac{1}{L_t(y_t)} \left[ \phi_{00}^t(y_t) - \phi_{01}^t(y_t) \right] \odot \left[ \Pr(S_{t-1} = 0|I_{t-1})p_{00}^t(1 - p_{00}^t) \right] \odot \left[ \Pr(S_{t-1} = 1|I_{t-1})p_{11}^t(1 - p_{11}^t) \right].$$

(15)

Here $\odot$ denotes the Hadamard product and $L_t(y_t)$ the (approximate) likelihood at time $t$.

It is common practice to scale the score before using it to drive a time-varying parameter, see Equation (6). Trivially, one could use no scaling (i.e. identity scaling). However scaling by a power of the (pseudo-)inverse Fischer matrix is often preferred. In particular, scaling with the square root (pseudo-)inverse Fischer matrix yields the attractive property of constant unit variances of the scaled scores. However, as the Fischer matrix is obtained from the expectation of the outer product of the score, this latter method presents a large computational burden here. We shall therefore take a more practical approach to scaling instead. To facilitate this process and with our empirical application in mind we proceed now with just a single time-varying transition probability.

When considering a single time-varying transition probability we propose to adjust its score in two ways before using it in update Equation (6). First, we correct for the effects of the link function chosen to ensure that $p_{kk}^t$ remains in the unit interval. That is, here the $p_{kk}^t(1 - p_{kk}^t)$ terms for $k \in \{0, 1\}$ found in the score, see Equation (15), are a remnant of the use of the logistic function to map the $f_{k,t}$ to the $p_{kk}^t$ by virtue of the chain-rule.

From a practical standpoint, with the transition probability empirically often being very close to either 0 or 1, the product $p_{kk}^t(1 - p_{kk}^t)$ is close to 0, very much dampening movement. Note also that it is straightforward to show that in the case of scaling with the square root (pseudo-)inverse Fischer matrix this part would also disappear, see Creal et al. (2013).

Second, we propose to consider a function of the resulting scaled score instead of using
it directly. The reason for this is motivated by empirical findings. Namely, it is found that in the DFMS model, even with square root (pseudo-)inverse Fischer scaling, the score can produce a large number of outliers. This in turn hampers its ability to drive the transition probability in a meaningful way. In a standard regression framework with a strictly positive regressor a straightforward remedy for this would be to consider the logarithm instead. However, as the score can be both positive and negative we propose the following intuitive extension:

\[ g(x) = \text{sign}(x) \log(1 + |x|). \] (16)

This monotonic function \( g(x) \) is antisymmetric around the origin, coinciding with the zero expectation of the score, that is close to the identity map for ‘small’ \(|x|\) and close to logarithm for ‘large’ \(|x|\). Note also that the expectation of \( g(s_t) \) may deviate from zero and even time-vary as the density of \( s_t \) changes. In practice, however, for our models this difference is small and of no significant consequence. To conclude, we propose the following process to drive time-variation of the transition probability \( p_{kk}^t \) both endogenously using the transformed score and exogenously using additional known explanatory variables as follows:

\[ f_{k,t+1} = w + a s_{t,k}^f + b f_{k,t} + \sum_{h=1}^{H} c_h x_{h,t}, \] (17)

whereby we have that \( s_{t,k}^f \) is given by

\[ s_{t,k}^f = g(\frac{\phi_t^{kk}(y_t) - \phi_t^{k(1-k)}(y_t)}{L_t(y_t)}) Pr(S_{t-1} = k|I_{t-1}), \] (18)

and \( w, a, b \) and the \( c_h \) are (scalar) parameters to be estimated.

### 2.3 Accelerated score-driven dynamics

In the accelerated score-driven framework the parameter that determines how much to update in the direction of the score is allowed to be time-varying by reapplying the score-driven
approach to this parameter. Specifically, Blasques et al. (2019) propose the following scheme for a time-varying parameter $f_{k,t}$:

$$f_{k,t+1} = w^{f_k} + a_{t+1} s^f_t + b^f_k f_{k,t},$$ (19)

with $s^f_t$ the scaled score with respect to $f_{k,t}$ and $w^{f_k}$ and $b^f_k$ parameters. Furthermore, it is then assumed that $a_{t+1}$ itself is driven by its score in similar fashion as follows:

$$a_{t+1} = a^a + a^o s^a_t + b^o a_t,$$ (20)

where $s^a_t$ denotes the scaled score with respect to $a_t$ and $w^a$, $a^o$ and $b^o$ parameters. Note that to maintain certain theoretical optimality properties one requires the non-negativity of $a_t$. Blasques et al. (2019) therefore propose adding a monotonic link function that maps to the positive domain. For simplicity no such link function is shown here. Additionally note that due to the chain rule, the score with respect to $a_t$ is the product of the unscaled score with respect to $f_{k,t}$ and its scaled score for the previous period $s^f_{t-1}$. If we reapply the same scaling factor as we did to obtain the scaled score $s^f_t$ from its raw score we have the following expression for $s^a_t$:

$$s^a_t = s^f_t s^f_{t-1}.$$ (21)

The scaled score with respect to $a_t$, i.e. $s^a_t$, is thus now simply the product of the scaled score with respect to $f_{k,t}$ and its one-period lag. Particularly, when scaling with the square root (pseudo-)inverse Fischer information of the score with respect to $f_{k,t}$, we have that the $s^f_t$ have a constant unit variance. Therefore the product of it with $s^f_{t-1}$ can be regarded as a proxy for its first-order autocorrelation. The intuition is simple: if recent scores align more a higher degree of movement is likely warranted. On the other hand if the scores misalign the information is more likely non-structural or noise, that should not be acted upon to a large degree. Here the addition of a monotonic link function that maps large negative values
of \( a_t \) to positive values near 0 would ensure that no (large) opposite steps are taken.

However, in view of the non-standard scaling of the previous section in the absence of a computationally feasible estimator of the Fischer information, we provide a simple and intuitive analog for our specification. That is, building upon the general idea of score-alignment to dictate the degree of the update, we propose the following approach. Roughly speaking the parameter \( a \) from Equation (17) will now become a time-varying parameter \( a_t \) and is made to vary smoothly between some parameters \( a^l > 0 \) and \( a^u > a^l \), to be estimated, by considering a feasible proxy of the correlation coefficient between \( s_{f_k}^t \) and \( s_{f_k}^{t-1} \). Specifically, we construct the parameter \( a_{t+1} \) as follows:

\[
a_{t+1} = a^l + a^u u_{t+1},
\]

(22)

whereby we have that \( u_{t+1} \) admits the following dynamics

\[
u_{t+1} = \delta u_t + (1 - \delta) \rho_t,
\]

(23)

for some parameter \( \delta \in (0, 1) \) to be estimated. Here, \( \rho_t \) denotes a feasible proxy of the correlation coefficient between \( s_{f_k}^t \) and \( s_{f_k}^{t-1} \) linearly mapped to the unit interval and is constructed as follows:

\[
\rho_t = \frac{s_{f_k}^t s_{f_k}^{t-1}}{(s_{f_k}^t)^2 + (s_{f_k}^{t-1})^2} + \frac{1}{2}.
\]

(24)

Note that the expression of \( \rho_t \) is similar to that of the (transformed) sample autocorrelation and may more generally be simply interpreted as an intuitive measure of alignment. For example if \( s_{f_k}^t \) and \( s_{f_k}^{t-1} \) have the same value, then \( \rho_t \) is maximal at 1. Conversely, if \( s_{f_k}^t \) and \( s_{f_k}^{t-1} \) have the same absolute value but are of opposite sign this would put \( \rho_t \) to 0.

The idea between this setup is as follows. First, by opting for integrated dynamics for \( u_t \) and having its innovation term \( \rho_t \) in the unit interval, \( u_t \) is also contained therein. This approach limits the number of additional parameters to be estimated and also does
not require us to consider an additional (non-linear) link function to ensure (at least) the positivity of $a_t$. Second, because here $s_{t}^{f_k}$ is a transform of the (scaled) score of $f_{k,t}$, it is not straightforward to apply their framework directly. Therefore by considering the first-order sample correlation of $s_{t}^{f_k}$ to help dictate its importance, we provide an intuitively clear alternative that adheres to the core rationale of their accelerated approach.

### 2.4 Estimation

As the DFMS model contains both unobserved regimes and an unobserved factor, estimation makes use of both the Hamilton filter and the Kalman filter. Additionally, to prevent a computationally intractable length of states, due to path-dependence, we employ a collapsing step as proposed by Kim (1994). The resulting prediction-update recursion is provided in Appendix A. To obtain the parameter estimates we maximize the approximated log-likelihood as given in Equation (12). To identify the factor, we fix the first factor loading i.e. $\lambda_1 = 1$. Additionally, to identify the regimes we set $\alpha_1 < 0$, such that $S_t = 1$ corresponds to a contraction and consequently $S_t = 0$ to an expansion. Optimization is done using standard quasi-Newton methods and standard errors are obtained from the inverse hessian. In terms of initialization, the contraction state probability at time $t = 0$ is set to 0, with virtually identical results obtained if this is treated as a parameter which is optimized alongside the other parameters. For the specifications that include a time-varying transition probability, the initial values are set to those obtained from the base model with all transition probabilities being constant. Optimization of these extended models is practically identical to that of the base model. Finally, for the initialization of the factor, a diffuse prior with zero mean is considered at time $t = 0$. 

13
3 Empirical application

3.1 Data

For our empirical application, we consider the four components of TCB’s CEI for the US economy: employees on nonfarm payrolls (EMP), the index of industrial production (IDP), manufacturing and trade sales (MAN) and personal income less transfer payments (INC). We analyze these variables at a monthly frequency from January 1959 until February 2020. The vintages for these series, which are used for the real-time exercise, were obtained from TCB and are supplemented with data from Jeremy Piger\textsuperscript{1} for the earlier vintages. A more elaborate discussion of the data and its handling, alongside a graphical illustration of the final vintage data can be found in Appendix B. The DFMS model and its extensions are estimated based on the logarithmic growth rates of these four indicators\textsuperscript{2}.

Finally, the interest rate spread (IRS), here constructed by subtracting the US Federal Funds (FF) rate from the 10-year US Treasury rate, is used as an exogenous variable. The variable is obtained from the Federal Reserve Bank of St. Louis database (FRED) and is used by TCB as one of the components in their Leading Economic Index (LEI). While a number of different exogenous variables may be considered, there is a rich history documenting the usefulness of incorporating yield curve information in this setting. Moreover, it is often used as a benchmark and still remains hard to beat in a variety of empirical applications, see again e.g. Rudebusch and Williams (2009) and Liu and Mönch (2016).

A graphical illustration of the IRS, an indicator for its negativity and how they relate

\textsuperscript{1}See https://pages.uoregon.edu/jpiger/research/published-papers and add ‘/raw-real-time-data.zip’ to download the data.

\textsuperscript{2}Due to the presence of some large outliers in INC, particularly in December 2012, an adjustment is made. Specifically, we first winsorize the (logarithmic) growth rates of INC on both sides at a 1 percent level and estimate the model parameters. Second, we return to the original data and search for a maximum in a sensible neighbourhood of this winsorized optimum. This approach therefore excludes outlier solutions and prevents structural breaks in the parameter estimates in real-time when an outlier enters the estimation window.
to the NBER recession periods, is provided in Figure 1. Here we observe that periods of distress in the bond market often precede recessions of the real economy.

Figure 1: US interest rate spread, an indicator for its negativity and the NBER recessions, January 1959 - February 2020.

![Image of Interest Rate Spread](image_url)

Note: The plot depicts the IRS constructed by subtracting the Federal Funds (FF) rate from the 10-year Treasury rate. The dark shaded areas in the negative domain reflect the periods of a negative IRS. Finally, the light shaded areas in the positive domain reflect the recession periods as determined by the NBER.

### 3.2 Model for empirical application

For the ex-post analysis using the full sample we consider a total of six (nested) specifications. The first is the base model with time-invariant transition probabilities, while the second and third introduce GAS and aGAS dynamics for $p_{01}$ respectively and are denoted similarly. The final three specification are similar to the first three but add an exogenous variable, in our case an indicator for a negative IRS, to bring about time-variation of $p_{01}$ and are denoted in order by Exo, GASX and aGASX. Finally, for the real-time analysis we consider only the base model and the aGASX specification.

For our empirical application we consider the two-regime one-factor DFMS model with AR(1) specifications for both the common factor as well as the idiosyncratic components. The choice for two regimes is motivated by our interest in dating business cycle turning points. Although more involved multi-state models, such as the three phase model by Sichel
(1994), can provide a better in-sample fit, it appears sensible here to consider just two regimes in view of the relative paucity of recessions found in the sample. The sizes of the eigenvalues of the correlation matrix of the growth rates of the four coincident indicators motivate the choice of a single factor. The choice of only a single lag for the factor and idiosyncratic components, similar to Chauvet (1998) for monthly data, is mainly to not over-complicate the already reasonably involved inference and estimation.

Because our Markov regimes determine only the mean growth rates and not the variance levels, we do not explicitly account for the Great Moderation here. This structural break in output volatility, see McConnell and Perez-Quiros (2000), poses mainly a significant problem, when not accounted for, in the presence of regime-dependent variances. While a seemingly simple solution would be to allow for a different variance level before and after 1984, this would complicate real-time analysis. This is because from a practical standpoint it is not directly clear at what point in time this phenomenon became known and should be introduced in the model.

In addition, it should be noted that while we assume constant regime-dependent growth rates, recent research suggests that this might be improved upon. Eo and Kim (2016) allow for evolving regime-specific mean growth rates and find a steady decrease in real gross domestic product (GDP) growth rates post 1945. More recently, Eo and Morley (2019) and Doz et al. (2020) find that the 2008 Great Recession appears to have a permanent effect on GDP growth. While for practical reasons not considered here, the addition of evolving regime-dependent mean growth rates presents an interesting avenue for future research.

Furthermore, as we are primarily interested in improving peak dating and find little evidence of strong relevant dynamics in $p^{11}$, we allow only $p^{00}$ to vary over time. In fact to facilitate this we do not estimate $p^{11}$ alongside the other parameters in the ML-procedure, but instead calculate it directly from the completed NBER recessions. Specifically, this boils down to dividing the total number of recession months minus the number of recessions by the total number of recession months. For the real-time exercise this means that the value for $p^{11}$
is updated the month after a trough announcement by the NBER. The rationale is as follows. By fixing $p_{11}$, the (conditional) duration of a recession is also fixed and guarantees that the recession regimes are sufficiently persistent. In practice, we find for the time-varying models that $p_{11}$ is reduced, if it is estimated in the ML-procedure, paired with large movement in $p_{00}$ before and during recession periods. This may lead to patterns that allow for expansion months during recession months with only moderately negative or even positive growth rates of the four coincident indicators. Fixing $p_{11}$ directly using past NBER data ensures no such undesirable patterns are admissible. The evolution of this dynamically estimated $p_{11}$ from NBER recession periods is given in Appendix Figure B2 and shows little variation over time. Unsurprisingly, the NBER based estimator falls well within any sensible confidence interval of the ML-estimate of $p_{11}$ in the base model when estimated alongside the other parameters.

For expositional purposes, the time-variation in $p_{00}$ is from this point on presented in terms of $p_{01} = 1 - p_{00}$, such that we may speak of the probability to switch to a recession. For the specification with only exogenous information to make $p_{01}$ vary over time, we will set $a$, the parameter of the score, to zero, but maintain the autoregressive structure. With regards to the aGAS specifications it is found for our empirical applications that the estimates of $\delta$ are statistically indistinguishable from 0. Consequently, $\delta$ in Equation (23) is set to 0. While the estimates of $a^t$ in the aGAS framework are also often found to be insignificant we keep it for the full-sample results, such that now all model specifications to be considered are nested, allowing for straightforward likelihood-ratio (LR) testing.

For our empirical analysis we will only make use of a single exogenous variable, namely an indicator for a negative IRS. Effectively, using this approach, the size of $p_{01}^t$ is made to be proportional to the duration of the inverted yield curve. This allows the probability to switch to a recession to accumulate during periods of a negative IRS and to decay in periods of a positive IRS, with rates of accumulation and decay determined by the parameter estimates. This specification is able to bridge the delay found between periods of turmoil in the bond market and a recession in the real economy in a simple manner. Due to the substantial
decrease in the size of the interest rates in recent years, we find the outlined approach to be the most straightforward way to make use of the interest rate spread. Results of several robustness checks, whereby the IRS is incorporated in a different manner can be found in Appendix Figure C1 and C2. This includes first a specification that considers the IRS as is, i.e. without any transformation. Second, a specification making using of a standardized IRS, which divides the spread by the sum of the 10-year Treasury rate and the Federal Funds rates, is considered. Third, a specification whereby we simply allow for two different transition probabilities depending on the sign of the IRS is considered. Appendix Figure C1 and C2 additionally contain findings for using the LEI as the exogenous variable, which also appears to be an effective choice. Note however that these results include all revisions of the LEI known at the final date. The reason for considering the IRS over the LEI is first because the former is found to be slightly more timely for dating peaks and has more synergy with the score-driven approach, which is naturally more coincident. Moreover, the IRS is a far simpler option as the composition and weighing of the LEI has changed significantly over the years. Considering more or a weighted combination of leading indicators to drive the transition probability is left for future research.

3.3 Full-sample estimation results

In this section we discuss the estimation results of the DFMS model and our extensions for the period January 1959 until February 2020 using the data vintages as released in March 2020. Hence, this includes all the revisions known at the final date. In Table 1, for brevity, only key parameter estimates for the six considered specifications are shown. The remaining parameter estimates can be found in Appendix Table C1.

In Table 1, we observe from the log likelihood and the Akaike information criterion (AIC) that the various specifications that allow for a time-varying \( p^{(1)} \) improve upon the base model to different degrees. This includes notable improvements from using the IRS as exogenous input and more modest improvements due to the endogenous information captured by the
Table 1: Key parameter estimates for the DFMS model with variants that consider a time-varying $p^{01}$.

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>GAS</th>
<th>aGAS</th>
<th>Exo</th>
<th>GASX</th>
<th>aGASX</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w$</td>
<td>-0.559</td>
<td>-0.761</td>
<td>-0.463</td>
<td>-0.553</td>
<td>-0.514</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.218)</td>
<td>(0.458)</td>
<td>(0.167)</td>
<td>(0.183)</td>
<td>(0.175)</td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>0.839</td>
<td>0.743</td>
<td>0.936</td>
<td>0.910</td>
<td>0.902</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.112)</td>
<td>(0.023)</td>
<td>(0.025)</td>
<td>(0.029)</td>
<td></td>
</tr>
<tr>
<td>$a$</td>
<td>1.207</td>
<td>0.979</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.399)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a^d$</td>
<td>0.000</td>
<td></td>
<td>0.220</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td></td>
<td>(0.859)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a^u$</td>
<td>2.602</td>
<td></td>
<td>1.261</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.996)</td>
<td></td>
<td>(1.295)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRS&lt;0</td>
<td></td>
<td>0.533</td>
<td>0.685</td>
<td>0.723</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.171)</td>
<td>(0.221)</td>
<td>(0.234)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p^{01}$</td>
<td>0.017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LogL</td>
<td>-1913.3</td>
<td>-1907.3</td>
<td>-1905.4</td>
<td>-1900.2</td>
<td>-1897.9</td>
<td>-1897.3</td>
</tr>
<tr>
<td>$k$</td>
<td>16</td>
<td>18</td>
<td>19</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>AIC</td>
<td>3858.5</td>
<td>3850.7</td>
<td>3848.8</td>
<td>3836.5</td>
<td>3833.9</td>
<td>3834.7</td>
</tr>
</tbody>
</table>

Note: This table presents the key parameter estimates for the base model and the extensions that allow for a time-varying transition probability $p^{01}$. Standard errors are displayed below the estimates in brackets and $k$ denotes the number of parameters used.

(accelerated) GAS dynamics. In terms of parameter estimates, we find that the parameter of the negative IRS indicator is positive and significant for all specifications. The positive sign is in accordance with economic theory, 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 $b$ suggest that $p^{01}$ is highly persistent.

Furthermore, LR tests indicate that the GAS dynamics are significant with or without 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 and are given by $LR = 11.87$ ($p = 0.003$) and $LR = 4.6179$ ($p = 0.032$) respectively. In addition, using the same approach, we find that the accelerated GAS dynamics are significant, at the 5 percent level, in the absence of the exogenous variable when testing against regular GAS dynamics ($LR = 3.909$, $p = 0.048$), but not in the presence of the exogenous variable when
testing against the GASX specification \((LR = 1.201, p = 0.27)\). Considering the AIC or the individual significance of the relevant parameters corroborates these findings. In particular, we find the GAS parameter \(a\) to be significant in both specifications, whereas the aGAS parameter \(a_u\) is significant only in the absence of the exogenous input. The lower bound \(a_l\) in the aGAS specifications is statistically indistinguishable from 0 in either case.

In order to evaluate the usefulness of the models, we consider the filtered state probabilities, i.e. the \(Pr(S_t = 1|I_t)\). A model is considered superior if these state probabilities more closely match the NBER recessions, whereby in particular an eye is kept on the timely identification of business cycle peaks. Figure 2 compares the filtered state probabilities of the six specifications with the NBER recession dates. For readability, we present the filtered state probabilities only for the base model, while for the other specifications we show the differences of their filtered state probabilities with the base model. Specifically, we subtract the filtered state probabilities of the base model from the filtered state probabilities of the specifications that allow for a time-varying \(p^{01}\).

We observe from Figure 2 that the base model by itself is quite successful in the identification of the business cycle regimes. That is, we observe that the filtered probabilities remain close to 0 during expansion periods and rapidly increase to levels close to 1 during recessions. We note however that the base model can be somewhat slow in picking up the business cycle peaks, most notably for the recessions starting in 1969 and 1973. For the other recessions this can entail that the filtered state probability can be quite low in the first month(s) of the recession, which in turn may lead to delays in the detection of the peaks of the cycle. In addition, we find that the model can be somewhat noisy at times, for example around 1964 and 1968 relatively large spikes are found. We also observe that the base model
Figure 2: Filtered state probabilities for the DFMS model with variants that consider a time-varying $p^{01}$.

Note: Filtered contraction state probabilities ($Pr(S_t = 1|I_t)$) for the base model are depicted in the top left panel. The remaining figures depict the filtered contraction state probabilities of the extensions minus those of the base model, denoted by $\Delta Pr(S_t = 1|I_t)$. Finally, the shaded areas reflect the recession periods as determined by the NBER.

can be slow to recognize a trough, particularly so for the 2001 recession$^3$.

$^3$For this recession, these findings could stem from what is known as a jobless recovery, whereby employment stayed low while output experienced steady growth. A possible explanation could be structural change whereby a reallocation of workers across industries may have hampered growth, see Groshen and Potter (2003). Chauvet and Hamilton (2006) note however that particularly the nonfarm payrolls employment series used here contains a very slow recovery for the 2001 recession and propose using a civilian employment metric instead. For practical considerations this alternative measure of employment is not considered here.
Most notably, we observe in Figure 2 that the addition of the IRS variable provides large improvements over the base model. That is, we observe higher recession state probabilities during recession periods and lower such probabilities during expansions when using the exogenous variable compared to the base model. In particular, the largest probability differences are found around the peak dates, precisely where they are deemed most important. When comparing the three specifications that make use of the IRS, we find that adding (a) GAS dynamics improves slightly upon the specification which makes use of just the exogenous variable by increasing the differences around most peaks. For the GAS and aGAS specifications in the absence of the exogenous variable, Figure 2 indicates that the GAS dynamics modestly improve upon the base model, whereby in turn the aGAS specification improves upon the GAS specification. Here we observe that the aGAS specification features larger contraction state probabilities around the start of recessions and is better at reducing false signals when compared to the GAS specification.

Table 2 contains an overview of the signaling performance for the NBER recessions by the considered model specifications, which confirms the graphical evidence in Figure 2. Here we consider the Area-Under-the-Receiver-Operating-Curve (AUC), a common measure for evaluating the quality of binary classification ability. In addition, we consider the average contraction state probability during recessions, non-recession periods and the first month of the recessions, denoted by $\mu^r$, $\mu^e$ and $\mu^p$ respectively. We observe that all extensions improve upon the signaling ability of the base model, with the largest improvements stemming from the addition of the IRS information, which, for example, leads to an increase of the overall AUC from 0.941 to 0.979. Furthermore, we find for peak dating specifically, that GAS dynamics are a useful addition either on its own, increasing $\mu_p$ from 0.267 to 0.290, or alongside the exogenous variable increasing $\mu_p$ from 0.528 to 0.594. When considering the ratio of $\mu^r$ and $\mu^e$, which can be seen as measure for contrast, we find that the aGAS specification achieves the highest value here. Indicating that in a purely score-driven framework the accelerated framework may bring improvements over simple score-driven dynamics.
Table 2: Signaling performance of the filtered state probabilities for the base DFMS models and extension that allow for a time-varying $p^{01}$.

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>GAS</th>
<th>aGAS</th>
<th>Exo</th>
<th>GASX</th>
<th>aGASX</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.941</td>
<td>0.950</td>
<td>0.954</td>
<td>0.979</td>
<td>0.978</td>
<td>0.977</td>
</tr>
<tr>
<td>AUC Pre</td>
<td>0.937</td>
<td>0.942</td>
<td>0.950</td>
<td>0.975</td>
<td>0.975</td>
<td>0.973</td>
</tr>
<tr>
<td>AUC Post</td>
<td>0.972</td>
<td>0.975</td>
<td>0.984</td>
<td>0.989</td>
<td>0.989</td>
<td>0.992</td>
</tr>
<tr>
<td>AUC Pred</td>
<td>0.894</td>
<td>0.916</td>
<td>0.920</td>
<td>0.961</td>
<td>0.965</td>
<td>0.961</td>
</tr>
<tr>
<td>$\mu^r$</td>
<td>0.647</td>
<td>0.685</td>
<td>0.689</td>
<td>0.779</td>
<td>0.802</td>
<td>0.799</td>
</tr>
<tr>
<td>$\mu^e$</td>
<td>0.066</td>
<td>0.063</td>
<td>0.052</td>
<td>0.063</td>
<td>0.065</td>
<td>0.064</td>
</tr>
<tr>
<td>$\mu^r/\mu^e$</td>
<td>9.731</td>
<td>10.832</td>
<td>13.319</td>
<td>12.337</td>
<td>12.419</td>
<td>12.402</td>
</tr>
<tr>
<td>$\mu^p$</td>
<td>0.267</td>
<td>0.290</td>
<td>0.288</td>
<td>0.528</td>
<td>0.594</td>
<td>0.596</td>
</tr>
</tbody>
</table>

Note: Signaling power of the filtered state probabilities for the NBER recessions is evaluated using the Area-Under-Receiver-Operating-Characteristic curve (AUC), whereby Pre and Post indicate the subperiod before and after 1984 respectively and Pred the AUC for the predicted state probabilities. Furthermore, we have that $\mu^{r(e)}$ represent the average filtered contraction state probability during NBER recession (expansion) periods and $\mu^p$ the average probability in the first recession month.

In addition, Table 2 contains the AUC of the predicted state probabilities. Here we observe qualitatively similar results, but overall larger improvements upon the base model. Indicating that with a minimum of a one month delay in data releases and if one is interested in making a prediction about the current state, then the addition of time-varying transition probability may be even more rewarding. To consider the effects of the Great moderation, see e.g Bernanke (2004) for an overview, we also consider the AUC of the filtered state probabilities before and after 1984. Here we find that the recessions post 1984 appear easier to date, for example the overall AUC is equal to 0.937 before 1984 and afterwards increases to 0.972.

By construction, the relative differences in filtered state probabilities and their signaling performance directly stem from the differences in the specification for the transition probability $p^{01}$. Therefore we now consider the time-evolution of this parameter in more detail. In Figure 3 the estimates of $p^{01}_t$ are compared to the constant estimate of the base model and the NBER recessions. Naturally, a model specification is considered superior if it has a heightened $p^{01}_t$ just before or at the start of a recession.

In Figure 3, we find that movement brought about by the GAS and aGAS specification
Figure 3: Transition probability $p_{it}^{01}$ for the DFMS model with variants that consider a time-varying $p^{01}$.

Note: This figure displays the evolution of transition probability $p_{it}^{01}$ over time for the extensions that allow for time-variation therein. The dotted line represents the constant $p^{01}$ estimated by the base model. Finally, the shaded areas reflect the recession periods as determined by the NBER.

appears sensible, but is mostly coincident. Therefore peaks are not signaled in advance, but rather the switches are made more extreme. This is not too surprising considering the fact that the a(GAS) specifications still only make use of coincident information. We do observe that the aGAS specification seems more successful in this regard. When comparing the aGAS specification to the GAS specification we observe that the peaks are amplified while the other signals are dampened. Indicating that, in a purely score-driven framework, the
idea of score-alignment for dictating the magnitude of the update appears promising in this application. For the three other specifications, which make use of the IRS, we observe large movements already before the start of recessions. Here adding GAS dynamics strengthens the movement of $p_{t}^{01}$ and also adds a small increase before the 1960 recession for which no yield curve signal is contained in the data. The addition of accelerated GAS dynamics to the exogenous information, further increases the movement of $p_{t}^{01}$ for the 1974, 1990, 2001 and 2008 recessions but reduces it somewhat for the 1960 and 1970 recessions, compared to the GAS specification. Therefore, as indicated by the estimates from Table 1 and the performance metrics from Table 2, accelerated dynamics do not appear to provide much increased benefits over simple GAS dynamics in the presence of the exogenous input.

We conclude that exogenous information, here an indicator for a negative IRS, can help improve the dating of recessions and their peaks in particular, when used to guide the transition probability to switch from an expansion to a contraction directly. Furthermore, more advanced methods using score information may be employed. In the absence of exogenous information, the GAS specification presents modest improvements over the base model, with accelerated GAS dynamics providing increased performance over simple GAS dynamics by amplifying correct signals and reducing erroneous ones. When combining exogenous and score information, we find that the GAS dynamics are still a useful addition, particularly for peak dating, but that additional accelerated dynamics do not provide much further increased benefit.

### 3.4 Real-time exercise

In the previous section we found that allowing for a time-varying transition probability to switch from an expansion to a recession may aid in the timely identification of business cycle peaks. However, estimating the model on the entire sample yields parameter estimates which would not be available in real-time. Furthermore, a plethora of revisions in the data themselves make it so that we have to rely on vintages for a fair assessment of the real-time
performance of our specifications. In this section we perform such a real-time exercise for the base DFMS model and the aGASX extension. For the aGASX specification we set $a'$, that is the lower bound in the smooth transition interval for the GAS parameter $a_t$, to 0 as it is mostly insignificant. Combined with again $\delta = 0$, we have now that $a_t = a^u p_{t-1}$, such that the magnitude of the update by score is made proportional to a proxy of its correlation with its predecessor. The choice for this aGASX specification over the GASX specification is twofold. First, this specification provides the lowest information criteria, as it produces a slightly better likelihood than the GASX model with now an equal number of parameters. Second, the accelerated model variant appears more stable for earlier vintages when not much data is available by reducing false signals due to its consideration of score-alignment. However, results are qualitatively similar without the accelerated component here.

Chauvet and Piger (2008) consider the data at the end of the month and restrict the sample at each point in time to the series for which the least amount of information is available. We instead proceed with the maximum amount of data at our disposal in the third week of each month, in line with the approach by Camacho et al. (2018). Specifically, this entails that at time $t$, we have EMP and IDP available up to and including time $t - 1$. For MAN and INC we are presented with additional delays in data publication, such that we have observations only up to and including time $t - 3$ and $t - 2$ respectively.

The results of the outlined real-time exercise, whereby the models are repeatedly re-estimated with all information available at each point in time, are supplied below. Table 3 presents a comparison of the signaling performance of the real-time filtered and predicted state probabilities of the considered specifications. We observe in Table 3 that the aGASX specification trumps the base model in all considered metrics, with higher contraction state probabilities during NBER recession periods ($\mu^r$) and lower such probabilities during NBER expansion phases ($\mu^e$). Most importantly, in the interest of dating peaks, we find that average state probability of the first recession month ($\mu^p$) is more than doubled(tripled) for the filtered(predicted) state probabilities. In addition, we find overall improved signaling
Table 3: Real-time signaling performance for the base DFMS model and the aGASX extension.

<table>
<thead>
<tr>
<th>Filtered</th>
<th>Base</th>
<th>aGASX</th>
<th>Predicted</th>
<th>Base</th>
<th>aGASX</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.940</td>
<td>0.973</td>
<td>AUC</td>
<td>0.894</td>
<td>0.960</td>
</tr>
<tr>
<td>$\mu_{t-1}^{r}$</td>
<td>0.751</td>
<td>0.864</td>
<td>$\mu_{t</td>
<td>t-1}^{r}$</td>
<td>0.620</td>
</tr>
<tr>
<td>$\mu_{t-1}^{c}$</td>
<td>0.123</td>
<td>0.113</td>
<td>$\mu_{t</td>
<td>t-1}^{c}$</td>
<td>0.140</td>
</tr>
<tr>
<td>$\mu_{t-1}^{c}/\mu_{t-1}^{r}$</td>
<td>6.083</td>
<td>7.625</td>
<td>$\mu_{t</td>
<td>t-1}^{c}/\mu_{t</td>
<td>t-1}^{r}$</td>
</tr>
<tr>
<td>$\mu^{p}$</td>
<td>0.185</td>
<td>0.491</td>
<td>$\mu^{p}$</td>
<td>0.116</td>
<td>0.369</td>
</tr>
</tbody>
</table>

Note: Signaling power of the probabilities for the NBER recessions is evaluated using the Area-Under-Receiver-Operating-Characteristic Curve (AUC). Furthermore, we have that $\mu_{i|j}^{r(e)}$ represent the average contraction state probability at time $i$ during NBER recession (expansion) periods with observations up to and including time $j$, which corresponds to the real-time estimation at time $j+1$. Finally, $\mu^{p}$ denotes the average state probability during the first month of the recessions within the evaluation sample.

performance of both the filtered and predicted contraction state probabilities for the NBER recessions in terms of the AUC. Moreover, these improvements of the aGASX method over the base methodology are found to be significant for both the filtered and predicted contraction states probabilities ($p = 0.035$ and $p = 0.005$ respectively) using the (one-sided) testing methodology of Hanley and McNeil (1983).

Figure 4 depicts the real-time filtered state probabilities for the base DFMS model and the aGASX specification. Additionally, the evolution of the switching probability $p_{t|t-1}^{01}$ is displayed. We observe in Figure 4 that the aGASX specification much better matches the NBER recession periods relative to the base model, in line with the statistics from Table 3. Moreover, the extension seems to be better able to identify business cycle peaks in four out of the five recessions that are contained in the evaluation sample. Most notably, it indicates that the 2008 recession could have been anticipated well in advance using the proposed framework. For the 1990 recession, the aGASX model initially appears to slightly dismiss the signal relative to the base model, although this difference is very minor. Interestingly, Stock and Watson found similar results using their leading indicator at the time when incorporating the IRS, see Hamilton (2011) for a discussion. Moreover, in Appendix Figure D1 a brief comparison between the performance of the model that uses just the exogenous variable
Figure 4: Real-time results for the base DFMS model and the aGASX extension.

Note: The top left and top right figures display the filtered state probabilities for the base and aGASX model respectively and the bottom left presents their difference (Base minus aGASX). Furthermore, the bottom right figure displays the evolution of the transition probability $p_0^1$. Finally, the shaded areas reflect the recession periods as determined by the NBER.

(i.e. the Exo specification) and the aGASX specification used here is given. Here we find that the small dip for the 1990 peak is found to be much worse without aGAS dynamics. Indicating that while the IRS here appears to provide the most benefit, the aGAS dynamics play a role as well, boosting the correct signals of the spread and dampening its mistakes. In sum, this underlines that a combination of exogenous and endogenous information for driving the transition probability may be particularly powerful. To investigate the effects of data revisions, Appendix Figure D2 contains the results obtained when making use of final vintage data. Although fit is improved somewhat for both specifications when using final vintage data, findings remain qualitatively unchanged.

Furthermore, as we observed the largest benefits for the aGASX specification in the two most recent recessions, we will further investigate the evolution of the state probabilities
there in more detail. Figure 5 depicts the evolution of the real-time smoothed contraction state probabilities for the recessions that began in 2001 and 2008. Appendix Figure D3 contains the full history of the filtered and smoothed state probabilities for both model specifications.

Figure 5: History of real-time smoothed state probabilities for the base DFMS model and the aGASX extension around the 2001 and 2008 recession.

Note: This figures depicts the real-time smoothed state probabilities around the 2001 (top) and 2008 (bottom) recessions for the base (left) and aGASX specification (right). The x-axis contains the estimation date and the y-axis the sample date for which a probability is constructed. Finally, the red dotted lines reflect the turning points as determined by the NBER.

In Figure 5 we observe that the aGASX specification produces much higher contraction state probabilities than the base model for the first months of the 2001 and 2008 recessions. As a result the aGASX specification is able to accurately provide a strong contraction signal as soon as these recessions enter the estimation window, whereas the base model requires
multiple months of additional information to do so. These differences are substantial and naturally of direct relevance for dating the peaks of the cycle. Also note that as more time passes the models appear to largely agree on past states. Appendix Figure D4 contains a plot similar to Figure 5 for the filtered state probabilities, here improvements are even more pronounced. All in all these findings indicate that the aGASX 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.

As our extension pertains to the transition probability to enter a recession, we shall focus our attention on peak dating. For this we use a straightforward 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’). Here, we first require \( Pr(S_t|I_T) < \tau \) and \( Pr(S_{t+k}|I_T) \geq \tau \), for \( k = 1,2,3 \) for some threshold \( 0.5 < \tau < 1 \) and where \( T \) is the maximum length of observations available at that time. Afterwards, peaks are identified by the point in time where the probability crosses a half, immediately preceding this period. Meaning that we find the smallest non-negative integer \( q \) such that \( Pr(S_{t-q-1}|I_T) < 0.5 \) and \( Pr(S_{t-q}|I_T) \geq 0.5 \). The peak date is subsequently identified to be \( t - q - 1 \) and as such refers to the final expansion month. Before another peak can be called we require that the state probability remain below \( \tau \) 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 own threshold. Here for simplicity, we established the first period \( t \), after a peak has been dated, for which \( Pr(S_t|I_T) \geq \tau \) and \( Pr(S_{t+k}|I_T) < \tau \), for \( k = 1,2,3 \) as a trough and as such characterizes the final recession month. By construction, we now have that a peak can only be called after a trough and vice versa. Because this procedure is done at each point in time with the available data (i.e. for each ‘column’), it might be that at later estimation dates different turning points are established than before. In Table 4 below the initial peaks obtained from the method outlined above for \( \tau = 0.65 \) and \( \tau = 0.8 \) are presented. To obtain the initial peaks (and troughs) we compare the most
recently suggested turning point at each estimation date with our most recently produced initial turning point. Here we select this newly suggested turning point only if it is a peak(trough) and our most recently established initial turning point is a trough(peak) and comes at a earlier date. By considering the recent state probabilities of the initial estimation date, December 1976, we determine that our evaluation window begins in an expansion and as such begin looking for a peak. Alternatively, dating may be made more sophisticated by adding duration, deepness and their combination, severity, into the dating decision, see Anas et al. (2007).

Table 4: Comparison peak dates from the base DFMS model and the aGASX extension with the NBER recessions.

<table>
<thead>
<tr>
<th>Peak date</th>
<th>Ann. date</th>
<th>Base</th>
<th>aGASX</th>
<th>NBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>Jan 1980</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>Jul 1981</td>
<td>-1</td>
<td>-2</td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
<td>Jul 1990</td>
<td>-4</td>
<td>-4</td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
<td>Mar 2001</td>
<td>-1</td>
<td>-4</td>
</tr>
<tr>
<td>-3</td>
<td>-5</td>
<td>Dec 2007</td>
<td>-7</td>
<td>-12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Peak date</th>
<th>Ann. date</th>
<th>Base</th>
<th>aGASX</th>
<th>NBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>Jan 1980</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>Jul 1981</td>
<td>-1</td>
<td>-2</td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
<td>Jul 1990</td>
<td>-4</td>
<td>-4</td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
<td>Mar 2001</td>
<td>-1</td>
<td>-4</td>
</tr>
<tr>
<td>-5</td>
<td>-6</td>
<td>Dec 2007</td>
<td>-5</td>
<td>-10</td>
</tr>
</tbody>
</table>

Note: This table contains the monthly differences in obtained peak dates of the base model and aGASX extension with the NBER database. 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.

In Table 4 we observe that the DFMS models are able to match or precede the announcement (ann.) dates made by the NBER, regularly by a substantial margin. However, while the NBER has not made any revisions since the inception of their dating method, the DFMS specifications makes adjustments for the 2008 peak date as more data becomes available. To see this, consider again Figure 5 above. Here we observe that our dating of the 2008 recession peak becomes closer to the date established by the NBER from early 2009 on. Although,
depending on the dating rule, the aGASX model still generally dates it somewhat earlier than the NBER. When comparing the base and aGASX specification, we note in particular a large increase in timeliness relative to the base model for the dating of the 2001 and 2008 business cycle peaks, in line with the probability differences found in Figure 5. More specifically, the aGASX specification is able to date the 2001 and 2008 peaks three and five months earlier respectively for both thresholds. For the most recent recession of 2008 the aGASX specification is even able to signal the peak a year before the announcement by the NBER using the threshold $\tau = 0.65$. For the first three recessions we note comparable performance of the base model and the extension, but note that both are able to match or precede the NBER announcements.

With regards to the troughs corresponding to the peaks of Table 4, we confirm the results of Chauvet and Piger (2008), in that the DFMS model in general is able to call troughs much earlier than the NBER. However, no noteworthy differences between the base DFMS model and the aGASX specification are found, with only a minor advantage for the 2001 trough for the base model for $\tau = 0.8$. This is somewhat unsurprising as the transition probability $p^{11}$ plays a much larger role here than $p_t^{01}$ and is set the same for both specifications. For this reason and for brevity the troughs corresponding to the peaks of Table 4 can be found Appendix Table D1. We do note that the troughs for the 1990 and especially the 2001 recessions are revised several times after their initial establishments and temporarily prolonged due to in part data revisions. This issue for the 2001 recession is also discussed in Chauvet and Piger (2008) and is presumably, as indicated before, related to a jobless recovery. Opting for a different measure of employment, such as civilian employment instead of the nonfarm payrolls used here and by TCB, is found by Doz et al. (2020) to remedy this dating delay.

We conclude that qualitatively our findings in real-time closely match those of our ex-post analysis. This entails that the aGASX specification that allows for a time-varying $p^{01}$ using both exogenous and endogenous information provides clear benefits over the base
DFMS model. This includes a significant increase in binary signaling ability of the filtered and predicted state probabilities for the NBER recessions, as measured by the AUC, and higher contraction state probabilities at the start of recessions. Additionally, we find that when converting the smoothed state probabilities to peaks that the DFMS specifications may be able to call turning points as soon or sooner than the NBER. Some adjustments of the turning point dates, when more data and revisions become available, are however noted. Moreover, the aGASX specification is able to date the two most recent peaks associated with the 2001 and 2008 recessions three and five months ahead of the base model respectively.

4 Concluding remarks

In this paper we have extended the DFMS model by Chauvet (1998), a framework that has been found useful in the empirical identification of the business cycle, by allowing for time-varying transition probabilities. Specifically, a score-driven approach is proposed to guide the time-evolution of the transition probabilities endogenously by building upon the techniques found in Bazzi et al. (2017). In addition, we consider a simplified analog of the recently developed aGAS framework by Blasques et al. (2019). This approach consider the recent alignment of scores in determining the magnitude of the update by the score. Furthermore, as suggested by Diebold et al. (1994), we propose letting the transition probabilities depend on relevant economic and financial variables. The resulting aGASX framework therefore makes use of both endogenous and exogenous information for driving transition probabilities within the DFMS framework.

An empirical application using the components of TCB’s CEI for the period 1959-2020 shows that the proposed method can improve the signaling power of the state probabilities for the NBER recessions. Specifically, we apply the aGASX framework to the probability to switch from an expansion to a recession state and make use of an indicator for a negative IRS as our exogenous input. We find that the improvements largely stem from the yield curve
information, with a more modest amplification role for the score-driven dynamics. Here the
autoregressive specification used for driving the transition probability is able to effectively
bridge the delay between periods of distress in the bond market and the (often) subsequent
recessions. A real-time exercise using appropriate vintages from December 1976 until March
2020 indicates that these findings can also be useful in a practical setting, with significantly
improved real-time signaling power for the NBER recessions. When converting smoothed
state probabilities to turning points, it is found that the aGASX framework can date the
recent 2001 and 2008 recession peaks 3 and 5 months faster than the base model.

A multitude of possibilities for further research exist. First, by finding exogenous inform-
ation to drive the transition probability to exit a recession in a meaningful way, trough
dating for recessions, particularly those associated with jobless recoveries, could possibly
be further improved. Second, as the timely identification of business cycle peaks crucially
depends here on the information found in the yield curve, future research could further
explore different methods to extract information from it and related variables in a more
efficient manner. Finally, besides interest rates, one could also consider measures related to
for example the stock-market, oil prices or consumer confidence among others to drive the
transition probability to switch to a recession period. The outlined framework presents a
straightforward approach in which these variables may be incorporated without the need of
specifying additional factors, Markov chains or feedback mechanisms.

References

Aastveit, Knut Are, Anne Sofie Jore, and Francesco Ravazzolo (2016). “Identification and
real-time forecasting of Norwegian business cycles”. In: International Journal of Fore-
casting 32:2, pp. 283–292.

Anas, Jacques et al. (2007). “A turning point chronology for the Euro-zone”. In: Growth and


Eo, Yunjong and James Morley (2019). “Why has the US economy stagnated since the Great Recession?” In: *Available at SSRN 3077864*.


Ng, Serena and Jonathan H Wright (2013). “Facts and challenges from the great recession for forecasting and macroeconomic modeling”. In: *Journal of Economic Literature* 51.4, pp. 1120–54.


A Prediction-update recursion

The Hamilton prediction step is given as

$$Pr(S_t = j, S_{t-1} = i | I_{t-1}) = p^{ij} Pr(S_{t-1} = i | I_{t-1})$$  \hspace{1cm} (25)

and the Kalman prediction steps can be written as

$$\zeta_{i|t-1}^{ij} = d_{s_t=j} + T\zeta_{i|t-1}^i,$$ \hspace{1cm} (26)

$$P_{i|t-1}^{ij} = TP_{i|t-1}^i T' + Q.$$ \hspace{1cm} (27)

Using the observation at time $t$, we can update our Markov state probability and the mean and covariance matrix of the state vector. The Hamilton update step is given by

$$Pr(S_t = j, S_{t-1} = i | I_t) = \left(\sum_{l,k \in \{0,1\}} p_{l|t}^{ik} Pr(S_{t-1} = l | I_{t-1})\phi_{l|t}^{ij}(y_t)\right) \phi_{j|t}^{ij}(y_t)$$ \hspace{1cm} (28)

and the Kalman update steps are given by

$$\zeta_{i|t}^{ij} = \zeta_{i|t-1}^{ij} + P_{i|t-1}^{ij} Z' (ZP_{i|t-1}^{ij} Z')^{-1} (y_t - Z\zeta_{i|t-1}^{ij}),$$ \hspace{1cm} (29)

$$P_{i|t}^{ij} = (I - P_{i|t-1}^{ij} Z' (ZP_{i|t-1}^{ij} Z')^{-1} Z) P_{i|t-1}^{ij}.$$ \hspace{1cm} (30)

To prevent an increasing length of path dependence, we collapse the states, a method first proposed in this context by Kim (1994). For the state probability this is relatively straightforward and given as

$$Pr(S_t = j | I_t) = \sum_{i \in \{0,1\}} Pr(S_t = j, S_{t-1} = i | I_t).$$ \hspace{1cm} (31)
The collapse step for the factor and its covariance are slightly more involved and given as

\[
\zeta^j_{t\mid t} = \frac{\sum_{i\in\{0,1\}} \Pr(S_t = j, S_{t-1} = i\mid I_t) \zeta_{ij}^{jt}}{\Pr(S_t = j\mid I_t)}, \tag{32}
\]

\[
P^j_{t\mid t} = \frac{\sum_{i\in\{0,1\}} \Pr(S_t = j, S_{t-1} = i\mid I_t) \left( \zeta_{ij}^{jt} + \left( \zeta_{ij}^{jt} - \zeta_{ij}^{jt} \right)(\zeta_{ij}^{jt} - \zeta_{ij}^{jt})' \right)}{\Pr(S_t = j\mid I_t)}. \tag{33}
\]

We now have all the components to begin the next iteration, i.e. start the prediction steps for \( t+1 \) and so on. To obtain the parameter estimates we maximize the associated approximated log-likelihood, which is here given as

\[
\log L \approx \sum_{t=1}^{T} \log \left[ \sum_{i,j\in\{0,1\}} p_{ij}^{jt} \Pr(S_{t-1} = i\mid I_{t-1}) \phi_t^{ij} (y_t) \right]. \tag{34}
\]

## B Data

### B.1 Employees on nonfarm payrolls and index of industrial production

The vintages for US total nonfarm payroll employment and the US industrial production index between December 1976 up to and including December 1988 are retrieved from the real-time dataset for macroeconomists from the Philadelphia Fed, see Croushore and Stark (2003). Data is reported after the monthly publication of the previous month’s employment, such that in month \( t \) we have vintages up to and including month \( t-1 \) for these variables. For January 1989 through March 2020 the vintages are obtained from TCB. Here data up to January 2001 are reported at start of the month before publication of the previous month’s employment, while the remainder of the data reports values later in the month after the publication of last month’s values. Because no new information becomes available between the end of each month and the beginning of its subsequent month for these variables, we shift the data for this period one month, such that it is available one month earlier. For
December 2001 two observations are reported, one in the first and one in the third week, such that the time-shift does result in missing data for this month.

**B.2 Manufacturing and trade sales**

Similar to Chauvet and Piger (2008), data for the vintages from December 1976 up to and including December 1988 are obtained from *Business Conditions Digest*. For the remaining period of January 1989 through March 2020, data is retrieved from TCB. For both datasets the reporting moment in each month $t$ is after the publication of the observation for month $t - 3$.

**B.3 Personal income less transfer payments**

Similar to Chauvet and Piger (2008), data for the vintages from November 1976 up to and including December 1988 are obtained from *Business Conditions Digest*. For the remaining period of January 1989 through March 2020, data is retrieved from TCB. For the first part of the data the reporting moment in month $t$ is just after the publication of the observation of month $t - 1$, while for the second part the reporting moment precedes this publication. Therefore, to synchronise the datasets, and to ensure we only include data up to the third week of each month, we lag the observations of the first dataset with one month. Now we have that for each month $t$ we have vintages up to and including month $t - 2$, with data available for the real-time analysis from December 1976 on. Furthermore, missing data are encountered for January 1997, where data starts in December 1992. The resulting missing log growth rates are filled with a constructed series using the same steps Chauvet and Piger (2008) employ to construct their personal income data after 1995. Specifically, nominal transfer payments are subtracted from nominal personal income and finally divided by the ratio of nominal to real disposable income. These components are collected from the ALFRED database, maintained by the Federal Reserve Bank of St. Louis. For nominal transfer payments, data from Economic Indicators, Business Statistics and the Survey of
Current Business are considered.

Figure B1: Final vintage US employees on nonfarm payrolls (EMP), the index of industrial production (IDP), manufacturing and trade sales (MAN) and personal income less transfer payments (INC), January 1959 - February 2020.

Note: Graphical illustration of the raw data, whereby the shaded areas reflect the recession periods as determined by the NBER.
Figure B2: Dynamic estimates of $p^{11}$ from completed NBER recessions.

Note: This figure depicts the evolution of the estimate of $p^{11}$ when estimated in an expanding window fashion from completed NBER recessions. The estimate is updated the month after a trough announcement. Finally, the shaded areas reflect the recession periods as determined by the NBER.

C Full-sample

Table C1: Remaining parameter estimates for the DFMS model with variants that consider a time-varying $p^{01}$.

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>GAS</th>
<th>aGAS</th>
<th>Exo</th>
<th>GASX</th>
<th>aGASX</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>0.094</td>
<td>0.091</td>
<td>0.091</td>
<td>0.093</td>
<td>0.093</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-0.097</td>
<td>-0.095</td>
<td>-0.111</td>
<td>-0.094</td>
<td>-0.089</td>
<td>-0.092</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.020)</td>
<td>(0.021)</td>
<td>(0.023)</td>
<td>(0.022)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.546</td>
<td>0.556</td>
<td>0.549</td>
<td>0.552</td>
<td>0.552</td>
<td>0.556</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.049)</td>
<td>(0.047)</td>
<td>(0.051)</td>
<td>(0.048)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>$\lambda_{IDP}$</td>
<td>2.298</td>
<td>2.310</td>
<td>2.305</td>
<td>2.302</td>
<td>2.305</td>
<td>2.308</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.115)</td>
<td>(0.116)</td>
<td>(0.115)</td>
<td>(0.115)</td>
<td>(0.116)</td>
</tr>
<tr>
<td>$\lambda_{MAN}$</td>
<td>1.907</td>
<td>1.914</td>
<td>1.915</td>
<td>1.909</td>
<td>1.910</td>
<td>1.910</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.109)</td>
<td>(0.110)</td>
<td>(0.109)</td>
<td>(0.109)</td>
<td>(0.109)</td>
</tr>
<tr>
<td>$\lambda_{INC}$</td>
<td>1.326</td>
<td>1.331</td>
<td>1.335</td>
<td>1.328</td>
<td>1.328</td>
<td>1.329</td>
</tr>
<tr>
<td></td>
<td>(0.075)</td>
<td>(0.075)</td>
<td>(0.075)</td>
<td>(0.075)</td>
<td>(0.075)</td>
<td>(0.075)</td>
</tr>
<tr>
<td>$\sigma^2_{\varepsilon_{EMP}}$</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>$\sigma^2_{\varepsilon_{IDP}}$</td>
<td>0.370</td>
<td>0.369</td>
<td>0.370</td>
<td>0.369</td>
<td>0.369</td>
<td>0.369</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>$\sigma^2_{\varepsilon_{MAN}}$</td>
<td>0.705</td>
<td>0.704</td>
<td>0.704</td>
<td>0.705</td>
<td>0.704</td>
<td>0.705</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>$\sigma^2_{\varepsilon_{INC}}$</td>
<td>0.268</td>
<td>0.268</td>
<td>0.268</td>
<td>0.268</td>
<td>0.268</td>
<td>0.268</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
<tr>
<td></td>
<td>Base</td>
<td>GAS</td>
<td>aGAS</td>
<td>Exo</td>
<td>GASX</td>
<td>aGASX</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>( \theta_{EMP} )</td>
<td>-0.486</td>
<td>-0.478</td>
<td>-0.475</td>
<td>-0.480</td>
<td>-0.479</td>
<td>-0.480</td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.068)</td>
<td>(0.066)</td>
<td>(0.070)</td>
<td>(0.069)</td>
<td>(0.069)</td>
</tr>
<tr>
<td>( \theta_{IDP} )</td>
<td>0.156</td>
<td>0.154</td>
<td>0.151</td>
<td>0.155</td>
<td>0.157</td>
<td>0.157</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>( \theta_{MAN} )</td>
<td>-0.234</td>
<td>-0.235</td>
<td>-0.236</td>
<td>-0.235</td>
<td>-0.235</td>
<td>-0.235</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.037)</td>
<td>(0.037)</td>
<td>(0.037)</td>
<td>(0.037)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>( \theta_{INC} )</td>
<td>-0.110</td>
<td>-0.112</td>
<td>-0.112</td>
<td>-0.111</td>
<td>-0.111</td>
<td>-0.111</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.037)</td>
<td>(0.037)</td>
<td>(0.038)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>( \sigma_{\eta}^2 )</td>
<td>0.015</td>
<td>0.014</td>
<td>0.014</td>
<td>0.014</td>
<td>0.015</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>LogL</td>
<td>-1913.3</td>
<td>-1907.3</td>
<td>-1905.4</td>
<td>-1900.2</td>
<td>-1897.9</td>
<td>-1897.3</td>
</tr>
<tr>
<td>( k )</td>
<td>16</td>
<td>18</td>
<td>19</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>AIC</td>
<td>3858.5</td>
<td>3850.7</td>
<td>3848.8</td>
<td>3836.5</td>
<td>3833.9</td>
<td>3834.7</td>
</tr>
</tbody>
</table>

Note: This table presents the leftover parameter estimates for the base model and the extensions that allow for a time-varying transition probability \( p^{01} \). Standard errors are displayed below the estimates in brackets and \( k \) denotes the number of parameters used.
Figure C1: Robustness exogenous variable choice, filtered state probabilities.

Note: Filtered contraction state probabilities \( Pr(S_t = 1 | I_t) \) for the base model are depicted in the top left panel. The remaining figures depict the filtered contraction state probabilities of the extensions minus those of the base model, denoted by \( \Delta Pr(S_t = 1 | I_t) \). Exo-(\cdot) denotes the autoregressive specification that only makes use of (\cdot) to drive \( p_t^{01} \). For the LEI monthly logarithmic growth rates are used. Dummy-IRS<0 denotes a specification that simply considers two different levels of \( p_t^{01} \) depending on the sign of the IRS. Finally, the shaded areas reflect the recession periods as determined by the NBER.
Figure C2: Robustness exogenous variable choice, transition probability $p_t^{01}$.

Note: This figure displays the evolution of transition probability $p_t^{01}$ over time for the extensions that allow for time-variation therein. The dotted line represents the constant $p_t^{01}$ estimated by the base model. Exo-(·) denotes the autoregressive specification that only makes use of (·) to drive $p_t^{01}$. For the LEI monthly logarithmic growth rates are used. Dummy-IRS<0 denotes a specification that simply considers two different levels of $p^{01}$ depending on the sign of the IRS. Finally, the shaded areas reflect the recession periods as determined by the NBER.
D Real-time

Figure D1: Comparison real-time filtered performance for the Exo and the aGASX specification.

![Exo and aGASX versus Base](image1.png)

Note: The left figure displays the differences of the real-time filtered contraction state probabilities of the Exo and aGASX specification with those of the base model (Base minus extension). Furthermore, the right figure displays the evolution of the transition probabilities $p_{01}^t$. Finally, the shaded areas reflect the recession periods as determined by the NBER.

Table D1: Comparison trough dates from the base DFMS model and the aGASX extension with the NBER recessions.

<table>
<thead>
<tr>
<th>Trough date</th>
<th>$\tau = 0.65$</th>
<th>Ann. date</th>
<th>$\tau = 0.8$</th>
<th>Ann. date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>aGASX</td>
<td>NBER</td>
<td>Base</td>
<td>aGASX</td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
<td>Nov 1982</td>
<td>-5</td>
<td>-5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>Mar 1991</td>
<td>-14</td>
<td>-14</td>
</tr>
<tr>
<td>21</td>
<td>20</td>
<td>Nov 2001</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>Jun 2009</td>
<td>-5</td>
<td>-5</td>
</tr>
</tbody>
</table>

Note: This table contains the monthly differences in obtained trough dates of the base model and aGASX extension with the NBER database. Troughs 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.
Figure D2: Comparison real-time filtered performance for the base and the aGASX specification using final vintages.

Note: The top left and top right figures display the filtered state probabilities for the base and aGASX model respectively and the bottom left presents their difference (Base minus aGASX) using final vintage data. Furthermore, the bottom right figure displays the corresponding evolution of the transition probability $p_{01}^t$. To prevent large movement of $p_{01}^t$ for some of the early vintages when not much data is available, we constrain the autoregressive parameter $b$ to lie between 0.8 and 1. For all points in time a maximum is found here that does not lie on the boundary. Finally, the shaded areas reflect the recession periods as determined by the NBER.
Figure D3: Complete history of real-time filtered and smoothed state probabilities for the base DFMS model and the aGASX extension.

Note: This figure depicts the filtered (top) and smoothed (bottom) state probabilities for the Base (left) and aGASX (right) specification. The x-axis contains the estimation date and the y-axis the sample date for which a probability is constructed.
Figure D4: History of real-time filtered state probabilities for the base DFMS model and the aGASX extension around the 2001 and 2008 recession.

Note: This figures depicts the real-time filtered state probabilities around the 2001 (top) and 2008 (bottom) recessions for the base (left) and aGASX specification (right). The $x$-axis contains the estimation date and the $y$-axis the sample date for which a probability is constructed. Finally, the red dotted lines reflect the turning points as determined by the NBER.