CESifo WORKING PAPERS

10366
2023

April 2023

Dynamic Mixture Vector Autoregressions with Score-Driven Weights

Alexander Georges Gretener, Matthias Neuenkirch, Dennis Umlandt

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10366*

Dynamic Mixture Vector Autoregressions with Score-Driven Weights

Abstract

We propose a novel dynamic mixture vector autoregressive (VAR) model in which time-varying mixture weights are driven by the predictive likelihood score. Intuitively, the state weight of the k-th component VAR model in the subsequent period is increased if the current observation is more likely to be drawn from this particular state. The model is not limited to a specific distributional assumption and allows for straight-forward likelihood-based estimation and inference. We conduct a Monte Carlo study and find that the score-driven mixture VAR model is able to adequately filter and predict the mixture dynamics from a variety of different data generating processes, which other observation-driven dynamic mixture VAR models cannot appropriately handle. Finally, we illustrate our approach by an application where we model the conditional joint distribution of economic and financial conditions and derive generalized impulse responses.

JEL-Codes: C320, C340, G170.

Keywords: dynamic mixture models, generalized autoregressive score models, macro-finance linkages, nonlinear VAR.

Alexander Georges Gretener
University of Kiel / Germany
gretener@qber.uni-kiel.de

Matthias Neuenkirch
Trier University / Germany
neuenkirch@uni-trier.de

Dennis Umlandt
University of Innsbruck / Austria
Dennis.Umlandt@uibk.ac.at

April 2, 2023
The authors gratefully acknowledge comments and suggestions by Jan Pablo Burgard and participants of the 15th International Conference on Computational and Financial Econometrics, the 29th Annual SNDE Symposium, the 28th International Conference on Computing in Economics and Finance, the IAAE Annual Conference 2022, the German Economic Association Annual Conference 2022, and the Statistical Week 2022.

1 Introduction

Regime-switching models have a long tradition in macroeconomics and finance. The most common approaches to capture regime-dependent non-linearities are the Markov-switching (MS) vector autoregressive (VAR) model (Krolzig 1997; based on the seminal work by Hamilton 1989, 1990), the threshold VAR model (Tsay 1998), and the smooth transition VAR model (Weise 1999; Camacho 2004). More recent papers propose mixture VAR (MVAR) models with K different components (or states), each being linear Gaussian VAR processes, and so-called mixture weights. These weights can be constant over time (Fong et al. 2007), time-varying based on a Markovian process (Kalliovirta et al. 2016), or time-varying using observable covariates (Burgard et al. 2019).

We propose a flexible alternative approach to construct MVAR models with time-varying mixture weights, which are driven by past observations of the endogenous variables. To infer the direction and intensity of the weight updating within our **S**core-driven **M**ixture **V**ector **A**utoregression (SMVAR), we follow the generalized autoregressive score (GAS)[1] approach developed by Creal et al. (2013) and Harvey (2013). They propose to update parameters of econometric models towards the direction of the gradient of the log likelihood function, the so-called score, evaluated at the current observation.[2] Consequently, time-varying parameters are pushed towards the direction of steepest ascent of the observational likelihood function as indicated by the gradient. Blasques et al. (2015) show that such an update reduces the Kullback–Leibler divergence (Kullback and Leibler 1951) between the true and model-implied conditional density in each step.

The derived updating scheme for the weight of each mixture component uses the scaled conditional density of the component model evaluated at the current observation. Intuitively, the procedure increases the weights of those mixture components that appear particularly

[1] Also referred to as score-driven (SD) model or dynamic conditional score (DCS) model.

[2] GAS models have been applied successfully in numerous applications in time series analysis and financial econometrics. See, for example, Harvey and Lange (2017) and Gorgi et al. (2019) for applications in volatility modeling or Oh and Patton (2018) and Bernardi and Catania (2019) for systemic risk applications.

likely given the current observation. Interestingly, this updating mechanism is similar to those for univariate score-driven Markov switching models (Bazzi et al. 2017) and more general dynamic adaptive mixture models (Catania 2021). Moreover, the (unconditional) scaled observation density is also used as a driving variable in the dynamic MVAR models of Kalliovirta et al. (2016) and Burgard et al. (2019). We justify their ad-hoc modeling choice within a GAS framework and provide a reasoning for using scaled component observation densities to capture mixture weight dynamics. The proposed SMVAR is more flexible than other dynamic MVAR approaches and not restricted to a Gaussian component model. In fact, parametric distributions for the component models can be freely specified. The (conditional) likelihood function of the SMVAR model can be evaluated directly and allows for straightforward likelihood-based estimation and inference.

We perform a Monte Carlo study to investigate the abilities and limitations of several dynamic MVAR models in recovering mixture dynamics from various data generating processes both, in-sample and out-of-sample. First, we examine the performance of the SMVAR in comparison to the MVAR assuming static mixture weights and the Gaussian MVAR (GMVAR) of Kalliovirta et al. (2016). We find that the SMVAR always outperforms the GMVAR in filtering and predicting the mixture weights. This is particularly pronounced in data generating processes for which the mixture weights are constant over an extended period of time, such as constant weights or structural breaks. One reason for this is that the GMVAR updates the weights always according to the most recent scaled observation density of a mixture component and does not encompass situations in which weights are constant. The updating mechanism in the GMVAR facilitates an elegant verification of theoretical properties such as stationarity, ergodicity, and a known (stationary) distribution. However, this comes at the cost of lower flexibility to capture weight processes that are not Markovian, as our simulation study shows. The SMVAR solves this issue and allows for more flexibility by including a constant and an autoregressive (AR) part in the mixture weight dynamics. Second, we compare the filtering and predictive ability of the SMVAR against the

Logit MVAR (LMVAR) of Burgard et al. (2019) which uses signals of the covariate factor to derive the mixture weight dynamics. Our results show that the SMVAR performs better than the LMVAR if the employed signal includes less than 50% of the *true* covariate information. Hence, the SMVAR is not only useful if the explaining factors for the dynamics are unknown, but also helpful as complement to an LMVAR analysis if the considered factors are uncertain.

As an empirical application, we model the joint distribution of the National Financial Conditions Index and real GDP growth for the period 1971q1−2020q4 using a two-state SMVAR. We show that the mixture weights identify a (tranquil) normal regime and a (volatile) economic and financial crisis regime. In particular, all NBER recessions are accompanied by large values of the crisis state weight with the mild and short recession of 2001 being the only exception. Moreover, the recessions are anticipated by a drastic change in the mixture weights (or indicated with only a short delay in the case of the 1990−1991 recession). On average, the economy is in the normal state during 75% of the time. Finally, our application also highlights the appealing feature of utilizing an AR term in the mixture weight dynamics. In this specific example, the persistence in the mixture weights amounts to 90% and leads to a smooth development of the weights over time.

An analysis of the model dynamics is carried out with generalized impulse response functions based on Koop et al. (1996). One particular advantage of the SMVAR is that we can also analyze the impulse responses on the mixture weights. In our application, we find that adverse shocks on economic growth and financial conditions significantly increase the conditional probability for switching into a crisis regime for several quarters with economic shocks having a larger impact. The increased switching probability enlarges the set of possible trajectories after a shock and considerably alters the impact of the growth shock on financial conditions as compared to the component-wise impulse response where a regime switch has been ruled out.

The remainder of the paper is organized as follows. Section 2 introduces and discusses a general framework that almost all (dynamic) mixture vector autoregressions have in common. Our proposed model — which includes local likelihood optimal mixture dynamics — is presented and discussed in Section 3. A Monte Carlo study evaluating the performance of the novel model is conducted in Section 4. The empirical application is presented in Section 5. Section 6 concludes.

2 (Dynamic) Mixture Vector Autoregressions

Let $(y_t)_{t=1}^{\infty}$ denote the d-dimensional time series that is is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and equipped with a filtration $\mathcal{F}_t = \sigma(\{y_t, \ldots, y_1\})$ representing the information set available at time t. We assume the existence of K different states of the economy with y_t being driven by a different VAR specification in each state. The state of the economy is represented by a sequence of random vectors $s_t = (s_{1,t}, \ldots, s_{k,t})^\top$ where in every period t either $s_{k,t} = 1$ (state k is active in period t) or $s_{k,t} = 0$ holds ($k = 1, \ldots, K$ and $\sum_{k=1}^{K} s_{k,t} = 1$). A general MVAR model with K mixture components is then given by

$$y_t = \sum_{k=1}^{K} s_{k,t} \left(\Phi_{k0} + \sum_{i=1}^{p_k} \Phi_{ki} y_{t-i} + \Omega_k^{\frac{1}{2}} \varepsilon_t \right) \tag{1}$$

where ϵ_t is a d-dimensional sequence of independently spherically distributed random vectors with an identity dispersion matrix and positive definite Ω_k. We additionally assume ϵ_t to be independent of y_s for $s < t$ and to be conditionally independent of s_t given $\mathcal{F}_{t-1}$.

The state vector process s_t is not observable in general. We model a probability distribution pinning down the so-called mixing weights $\alpha_{i,t} = \mathbb{P}[s_{i,t} = 1 | \mathcal{F}_{t-1}]$ to derive a probability density for the time series of interest y_t. Given a particular specification of these weights,

the conditional probability density function (pdf) can be obtained by

$$p(y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_{k,t} p_k(y_t|\mathcal{F}_{t-1}) \tag{2}$$

where $p_k(y_t|\mathcal{F}_{t-1})$ is the conditional pdf of the k-th VAR model component.

To this point, the empirical framework nests a variety of mixture VAR models. First, there is the special case of constant mixture weights, that is, $\alpha_{i,t} \equiv \alpha_i$. The properties and estimation approaches for this MVAR model are discussed in Fong et al. (2007). Second, the popular Markov-Switching VAR (MSVAR) model (Krolzig 1997) is also a special case of the MVAR framework above, in which the mixture weights are given by $\alpha_{i,t} = \mathbb{P}[s_{i,t} = 1|s_{1,t-1}, s_{2,t-1}, \ldots, s_{K,t-1}]$. Third, the dynamic MVAR models of Kalliovirta et al. (2016) and Burgard et al. (2019) — that differ in the specification of the dynamic mixture weights $\alpha_{i,t}$ — are encompassed as well. Our novel specification also builds on the common framework above. However, it is much more flexible than the other dynamic specifications and optimal with regard to a local likelihood criterion.

3 Score-Driven Mixture Vector Autoregressions

We introduce the score-driven mixture VAR for a general number of mixture components K and discuss its properties. In particular, we compare the SMVAR to other dynamic MVAR specification and present a likelihood-based estimation strategy.

3.1 Score-Driven Mixture Weights

The updating scheme for the mixtures must ensure that weights sum up to one in each period, that is, the weights should result from the K-dimensional probability simplex $\mathbb{S}^K = \left\{\alpha \in \mathbb{R}^K | \sum_{i=k}^{K} \alpha_k = 1\right\}$. We achieve this by using a mapping $h : \mathbb{R}^{K-1} \to \mathbb{S}^K$ with $h(\tilde{\alpha}_t) = \alpha_t$ for $\tilde{\alpha}_t \in \mathbb{R}^{K-1}$. This ensures that the time-varying mixture weights α_t sum up to one,

while the dynamics of $\tilde{\alpha}_t$ are modeled in the unrestricted domain space. We now define the parameter updating, in line with Creal et al. (2013) and Harvey (2013), as

$$\tilde{\alpha}_t = \omega + \sum_{i=1}^{p} A_i S_t^{-1} \nabla_{t-i} + \sum_{i=1}^{q} B_i \tilde{\alpha}_{t-i} \tag{3}$$

with[3]

$$\nabla_t = \frac{\partial \ln p(y_t|\mathcal{F}_{t-1})}{\partial \tilde{\alpha}_t} = \mathcal{J}_h(\tilde{\alpha}_t) \cdot \begin{pmatrix} \frac{p_1(y_t|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})} \\ \vdots \\ \frac{p_K(y_t|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})} \end{pmatrix} \tag{4}$$

where $\mathcal{J}_h(\alpha_t)$ is the Jacobian of the mapping h, ω is a $(K-1)$-dimensional parameter vector, and A_i, B_i are $(K-1) \times (K-1)$-dimensional parameter matrices. Intuitively, this updating increases the future weights of the components with the highest observation density in the current period. S_t is a matrix that scales the impact of the observations on the parameter updating. A common choice is the Fisher information matrix $\mathcal{I}_t = \mathbb{E}\left(\nabla_t \nabla_t^\top | \mathcal{F}_{t-1}\right)$ or the Cholesky factor thereof to relate the scaling to the variance of the likelihood score.[4] However, the Fisher information of observation densities cannot be derived in closed-from in our particular mixture model. Therefore, for simplicity, we continue with another frequent choice and set S_t equal to the identity matrix, that is, $S_t = I$.[5]

The mixture weight updating derived in Eqs. (3) and (4) is particularly appealing since it relates the updating direction and intensity for the next period to the relative observation

[3] A derivation of the score ∇_t is provided in A.1.

[4] Using the Cholesky factor is particularly attractive as the parameter process exhibits unit variances in this case.

[5] The literature on GAS Models acknowledges that the choice of the scaling does not crucially affect the model performance in many applications. Accordingly, we find that using a numerically computed Fisher information via $\mathcal{I}_t = \mathcal{J}_h(\alpha_t) H_t \mathcal{J}_h(\alpha_t)^\top$ with

$$(H_t)_{i,j} = \int \frac{p_i(y_t|\mathcal{F}_{t-1}) p_j(y_t|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})} dy_t \tag{5}$$

does not crucially improve the model performance.

density of the component to be updated in the current period. Put differently, if the weighted observation density of the k-th component is high (low) in the current period compared to the overall observation density of all states, the mixture weight of state k will be increased (decreased) in the following period. Another attractive feature of the score-driven mixture weights updating is that it is invariant with respect to the component distribution models. Hence, the updating is valid not only for Gaussian component models (as assumed for many dynamic mixture VAR specifications), but also for other distributions.

Next, we have to specify the particular parameter updating to implement the SMVAR model. One possible choice is the logistic transformation where the mapping is defined as

$$h_k(\tilde{\alpha}) = \begin{cases} \dfrac{\exp(\tilde{\alpha}_k)}{1+\sum_{i=1}^{K-1}\exp(\tilde{\alpha}_i)}, & k = 1,\dots,K-1 \\ 1-\sum_{j=1}^{K-1}\dfrac{\exp(\tilde{\alpha}_j)}{1+\sum_{i=1}^{K-1}\exp(\tilde{\alpha}_i)}, & k = K \end{cases} \tag{6}$$

with the Jacobian given by[6]

$$(\mathcal{J}_h(\tilde{\alpha}))_{k,l} = \begin{cases} \dfrac{\exp(\tilde{\alpha}_k)\left(1+\sum_{i=1}^{K-1}\exp(\tilde{\alpha}_i)\right)-\exp(2\tilde{\alpha}_k)}{\left(1+\sum_{i=1}^{K-1}\exp(\tilde{\alpha}_i)\right)^2}, & k = l, k \neq K \\ \dfrac{-\exp(\tilde{\alpha}_k)\exp(\tilde{\alpha}_l)}{\left(1+\sum_{i=1}^{K-1}\exp(\tilde{\alpha}_i)\right)^2}, & k \neq l, k \neq K \\ \dfrac{-\exp(\tilde{\alpha}_l)}{\left(1+\sum_{i=1}^{K-1}\exp(\tilde{\alpha}_i)\right)^2}, & k = K \end{cases} . \tag{7}$$

This logistic transformation is also used by Bazzi et al. (2017) for creating score-driven dynamics of the conditional probability in univariate MS models. There are alternative transformations that could be considered, for example, those used by Catania (2021) for the more general dynamic adaptive mixture class. However, we did not encounter any striking

[6] A derivation of the Jacobian is provided in Appendix A.2.

improvement in our VAR context and, consequently, decided to keep the rather simplistic logistic transformation.

3.2 Two-Component Models

For simplicity, we present and discuss the properties of the SMVAR model only for the two-regime case. The model outlined in the following is also used in the simulation study in Section 4 and the empirical application in Section 5. In a two-regime model, the observation density simplifies to

$$p(y_t|\mathcal{F}_{t-1}) = \frac{\exp(\tilde{\alpha}_t)}{1+\exp(\tilde{\alpha}_t)} p_1(y_t|\mathcal{F}_{t-1}) + \frac{1}{1+\exp(\tilde{\alpha}_t)} p_2(y_t|\mathcal{F}_{t-1}). \tag{8}$$

Note that one internal latent process $\tilde{\alpha}_t$ suffices to describe the dynamic weights $\alpha_{1,t} = \exp(\tilde{\alpha}_t)/[1+\exp(\tilde{\alpha}_t)]$ and $\alpha_{2,t} = 1/[1+\exp(\tilde{\alpha}_t)]$ of the states $k=1$ and $k=2$, respectively. Similarly, the Jacobian of the logistic transformation can be expressed as

$$\mathcal{J}_h(\tilde{\alpha}) = \left(\frac{\exp(\tilde{\alpha})}{(1+\exp(\tilde{\alpha}))^2}, \ \frac{-\exp(\tilde{\alpha})}{(1+\exp(\tilde{\alpha}))^2} \right) \tag{9}$$

for $K=2$. Hence, we can derive the mixture weight updating as

$$\tilde{\alpha}_{t+1} = \omega + a \frac{\exp(\tilde{\alpha}_t)}{(1+\exp(\tilde{\alpha}_t))^2} \frac{p_1(y_t|\mathcal{F}_{t-1}) - p_2(y_t|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})} + b\tilde{\alpha}_t. \tag{10}$$

The updating scheme in Eq. (10) is highly intuitive. An update of the mixture weights is induced by a non-zero value of the scaled difference of the two state observation densities $[p_1(y_t|\mathcal{F}_{t-1}) - p_2(y_t|\mathcal{F}_{t-1})]/p(y_t|\mathcal{F}_{t-1})$. If the current observation y_t is more likely to be drawn from the first component VAR model (indicated by $p_1(y_t|\mathcal{F}_{t-1}) > p_2(y_t|\mathcal{F}_{t-1})$ in the numerator), the latent variable $\tilde{\alpha}_{t+1}$ is increased. This induces an increase in the weight of state 1 $\alpha_{1,t+1}$ in the following period and a decrease in the weight of state 2 $\alpha_{2,t+1}$. The scaling term $\exp(\tilde{\alpha})/[1+\exp(\tilde{\alpha})]^2$ results from the chosen weight transformation h.

3.3 Differences to Other Dynamic Mixture VAR Models

We briefly discuss other dynamic MVAR models and relate them to the SMVAR described in the previous subsections.

3.3.1 Gaussian Mixture Vector Autoregression

The Gaussian Mixture Vector Autoregressive model of Kalliovirta et al. (2016) employs the same baseline model as described in Section 2, but with Gaussian innovations and equal lag lengths for the VAR components representing the different states, hence $p_k = \overline{p}$ for $k = 1, \ldots, K$. For the weight updating, they stack $\overline{p}$ lags of the time series model in a $d\overline{p}$-dimensional vector $\mathbf{y}_t = vec(y_t, y_{t-1}, \ldots, y_{t-\overline{p}+1})^\top$. Hence, the regime component k features the common density

$$p_k(\mathbf{y}_t) = (2\pi)^{-d\overline{p}/2} \det(\Sigma_k)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}_t - \mathbf{1}_{\overline{p}} \otimes \mu_k)^\top \Sigma_k^{-1}(\mathbf{y}_t - \mathbf{1}_{\overline{p}} \otimes \mu_k)\right\} \tag{11}$$

with unconditional mean μ_k and covariance matrix Σ_k, which are functions of parameter matrices $\Phi_0, \Phi_1, \ldots, \Phi_{\overline{p}}$. The weight updating sets the weight as the ratio of the observation density p_k of each separate regime over the overall observation density p of the mixture model, that is,

$$\alpha_{k,t} = \alpha_k \frac{p_k(\mathbf{y}_{t-1})}{p(\mathbf{y}_{t-1})} \tag{12}$$

where $\alpha = (\alpha_1, \ldots, \alpha_K)^\top \in \mathbb{R}^K$ are unknown parameters. This parameterization also ensures that the dynamic weights stay within the unit simplex, that is, $\sum_{k=1}^K \alpha_{k,t} = 1$. Kalliovirta et al. (2016) show that the resulting process is stationary (if the component VARs are stationary) as well as ergodic and establish asymptotic normality of the maximum likelihood (ML) estimator.

The forcing variables for the dynamic weighting in the GMVAR in Eq. (12) and the SMVAR in Eq. (4) are conceptually similar. Both approaches infer information from ratios between component model densities and the density of the overall mixture model. While the SMVAR relies on conditional observation densities, the common density of p recent observations pins down the dynamic weights in the GMVAR. However, the SMVAR offers more flexibility in describing the data. First, our model allows specifying component models with different lag lengths. Second, the SMVAR model allows reducing or increasing the number of observations considered for the weight updating, whereas the GMVAR always uses the p most recent observations, as determined by the component VARs' lag length. Third, our model explicitly allows for non-Gaussian innovations in the component VARs. Fourth, another crucial difference is the AR term in the score-driven updating in Eq. (4). This allows modeling persistence in the mixture weights and typically reduces the number of lags required in the updating scheme.

3.3.2 Logit Mixture Vector Autoregression

The Logit Mixture Vector Autoregressive model of Burgard et al. (2019) also employs the same baseline model as described in Section 2, but defines the conditional mixture weights using a logit model. Hence, the conditional mixture weights are given by

$$\alpha_{k,t} = \frac{\exp\{\gamma_k x_{t-1}\}}{\sum_{j=1}^{K} \exp\{\gamma_j x_{t-1}\}} \tag{13}$$

with x_{t-1} being a vector of $\mathcal{F}_{t-1}$-measurable predictors and γ_k being a regime-dependent parameter vector. The set of predictors may include external variables and lags of the endogenous variables of the VAR. In addition, Burgard et al. (2019) and Bennani et al. (2022) use the component ratio from Eq. (12) of the GMVAR as predictor of the weight dynamics. Hence, the component ratio driving the SMVAR mixture update is also used as driving variable. Both, the LMVAR and the SMVAR, use a logistic transformation for mapping

weights onto the unit interval. However, the SMVAR maps an unrestricted internal latent process, whereas the LMVAR directly models the weights as dependent variables within an internal multinomial logit regression. The main advantage of the LMVAR is that external information can be incorporated into the mixture dynamics. Yet, this comes at the cost of potential misspecification. In contrast, the SMVAR does not require pre-specified predictors and the simulation study in Section 4 shows that it tracks mixture dynamics quite adequately without using external predictor information, thereby avoiding specification errors.

3.4 Estimation and Inference

We estimate the vector $\theta_k = (\Phi_0', vec(\Phi_{k1})', \ldots, vec(\Phi_{k,p_k})', vech(\Omega_k)')$ containing the VAR parameters for each component submodel. Furthermore, we estimate the parameters characterizing the mixture weight updating $\theta_s = (\omega', vec(A_1)', \ldots, vec(A_p)', vec(B_1)', \ldots, vec(B_p)')$ as well as possible additional parameters characterizing the distribution of φ collected in vector θ. In summary, this yields the complete parameter vector $\vartheta = (\theta, \theta_1, \ldots, \theta_k, \theta_s)$ with at least $(1+d+d(d+1)/2)K+(p+q)K^2+\sum_{k=1}^{K} p_k d^2$ entries. The conditional (log) likelihood function of the SMVAR model can be evaluated directly with

$$\ln \mathcal{L} = \ln \prod_{t=1}^{T} p(y_t|\mathcal{F}_{t-1}) \tag{14}$$

$$= \sum_{t=1}^{T} \ln \sum_{k=1}^{K} \alpha_{k,t} p_k(y_t|\mathcal{F}_{t-1}) \tag{15}$$

$$= \sum_{t=1}^{T} \ln \sum_{k=1}^{K} \alpha_{k,t} \varphi \left(\Omega_k^{-\frac{1}{2}} (y_t - \Phi_{k0} - \Phi_{k1} y_{t-1} - \ldots - \Phi_{kp_k} y_{t-p_k}) \right) \tag{16}$$

where $\alpha_{k,t}$ is defined as in Eq. (3).

Inference is conducted in the standard fashion for ML estimators as suggested by Creal et al. (2013) for GAS models. If ϑ stacks all the static parameters of the model, standard asymptotic theory for ML estimators would suggest that, under some regularity conditions,

the following holds:

$$\sqrt{T}\left(\hat{\vartheta}-\vartheta\right)\xrightarrow{d}\mathcal{N}\left(0,\mathcal{I}^{-1}(\vartheta)\right) \tag{17}$$

with the Fisher information matrix $\mathcal{I}(\vartheta):=-\mathbb{E}\left(\partial^2 l_t/\partial\vartheta\partial\vartheta^\top\right)$ where l_t is the log-likelihood contribution of the i-th observation evaluated at ϑ.

Likelihood functions of mixture models often suffer from having many plausible local optima. It is therefore advisable to use various starting values for the optimization to obtain what is likely to be the global maximum. For this reason, we employ the OQNLP multistart heuristic algorithm of Ugray et al. (2007). For deriving standard errors, we numerically compute the Fisher information matrix with a finite difference scheme.

3.5 Impulse Responses

A popular VAR-based tool for investigating the impact of shocks in empirical macroeconomics is the impulse response analysis. Given a fitted and identified VAR model as well as initial values, responses of a shock with known size and sign can be calculated in a straightforward manner. However, this conventional approach is not applicable for the SMVAR due to its crucial non-linearities and incompletely known initial values. Instead, we follow Koop et al. (1996) and compute generalized impulse response functions (GIRFs) given by

$$GI\left(n,\nu_t,\mathcal{F}_{t-1}\right)=\mathbb{E}\left(y_{t+n}|\nu_t,\mathcal{F}_{t-1}\right)-\mathbb{E}\left(y_{t+n}|\mathcal{F}_{t-1}\right) \tag{18}$$

where n is the number of periods ahead and ν_t a shock that occurs in period t.[7] Hence, Eq. (18) defines the impulse response as difference between the n-period ahead prediction given the shock and the corresponding prediction in absence of the shock. For a meaningful economic interpretation, the shock ν_t should be identified. This can be achieved by defining

[7] Virolainen (2020) uses the same concept to derive impulse responses for a structural version of the GMVAR model.

Algorithm 1: Generalized Impulse Responses for the SMVAR

1.) Estimate the model parameters of the SMVAR and pick an information set $\mathcal{F}_{t-1}$ (usually represented by a sequence $(y_i)_{i=0}^{t-1}$).

2.) **for** *r = 1,...,R* **do**

2.a) Draw $N+1$ random shocks $u_t^{(r)}, u_{t+1}^{(r)} \dots u_{t+N}^{(r)}$ from $\mathcal{N}(0, I_d)$.

2.b) Compute the reduced form shocks $\varepsilon_{t+n}^{(r)} = L u_{t+n}^{(r)}$ for $n = 0, \dots, N$.

2.c) Iterate the SMVAR with the reduced form shocks $\left(\varepsilon_{t+n}^{(r)}\right)_{n=0}^{N}$ to compute $\left(y_{t+n}^{(r)}(\mathcal{F}_{t-1})\right)_{n=0}^{N}$.

2.d) Iterate the SMVAR with the time t shock ν_t and the reduced form shocks $\left(\varepsilon_{t+n}^{(r)}\right)_{n=1}^{N}$ to compute $\left(y_{t+n}^{(r)}(\nu_t, \mathcal{F}_{t-1})\right)_{n=0}^{N}$.

2.e) Calculate

$$GI^{(r)}(n, \nu_t, \mathcal{F}_t) = y_{t+n}^{(r)}(\nu_t, \mathcal{F}_{t-1}) - y_{t+n}^{(r)}(\mathcal{F}_{t-1})$$

for $n = 0, \dots, N$.

end

3.) Form averages over the Monte Carlo replications:

$$\widehat{GI}(n, \nu_t, \mathcal{F}_t) = \frac{1}{R} \sum_{r=1}^{R} GI^{(r)}(n, \nu_t, \mathcal{F}_t)$$

the structural shocks u_t such that $\varepsilon_t = L u_t$ for a d-dimension matrix L. We then assume that the shock ν_t affects the VAR system via $L\nu_t$.

We follow a Monte Carlo approach to derive the generalized impulse responses to the shock ν_t in (18). The main idea is to simulate a large number of random sequences $(y_{t+n}(\mathcal{F}_{t-1}))_{n=0}^{N}$ where the shocks $(u_{t+n})_{n=0}^{N}$ are randomly drawn. Furthermore, a second series $(y_{t+n}(\nu_t, \mathcal{F}_{t-1}))_{n=0}^{N}$ is constructed for which the same random shocks are used but additionally $u_t = \nu_t$ is imposed. Then, we can use the means of these artificially drawn series to approximate the expectations in (18) and compute GI by taking differences. The procedure is described in more detail in Algorithm 1.

To avoid conditioning on an information set, another Monte Carlo loop can be added around Algorithm 1 in which $\widehat{GI}$ is integrated over draws from the historical sequence $(y_i)_{i=0}^{t-1}$.

Moreover, the algorithm can be easily adjusted for different distributions that are possible in the SMVAR framework by replacing the innovation distribution in step 2.a).

4 Monte Carlo Study

We conduct a simulation study to investigate the performance of the SMVAR model in comparison to alternative (dynamic) MVAR models in the literature.

4.1 Data-Generating Processes

We base the Monte Carlo study on simulated time series with dimension $d = 2$ that are affected by two regimes or states ($K = 2$). The two regimes feature conditional normal distributions with the following parameter configurations:

$$\begin{aligned} y_t =& s_{1,t} \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0.5 & 0.2 \\ 0.2 & 0.5 \end{pmatrix} y_{t-1} + \begin{pmatrix} 0.5 & 0 \\ 0.2 & 0.5 \end{pmatrix} \epsilon_t \right) \\ &+ s_{2,t} \left(\begin{pmatrix} -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.6 & 0 \\ 0.4 & 0.6 \end{pmatrix} y_{t-1} + \begin{pmatrix} 0.2 & 0 \\ 0.8 & 0.2 \end{pmatrix} \epsilon_t \right). \end{aligned} \tag{19}$$

We consider several dynamic processes to describe the evolution of the mixture weight in the data-generating process (DGP). These include deterministic paths, such as a time-constant weight, a structural break, and a cycle. In addition, we simulate a Markovian regime-switching model that depends on prior realizations of the latent variable s_t and a logit transform of an autoregressive process of order one. The processes in the simulation study are summarized in Table 1.

We simulate the DGP with different mixture weight dynamics for time series lengths $T = 200, 500, 1000$ and then estimate the parameter vector ϑ for a SMVAR with lags $p = q = 1$ and $K = 2$ using likelihood optimization. In addition, we fit an MVAR, an MSVAR, a GMVAR, and an LMVAR as benchmark to evaluate the performance of our SMVAR model.

Table 1: Mixture Processes for Simulation Study

Model		**Mixture Weights** (α_t)
I	Constant	0.8
II	Break	$0.8 \cdot \mathbb{1}\left(t < \frac{T}{2}\right) + 0.2 \cdot \mathbb{1}\left(t \geq \frac{T}{2}\right)$
III	Cycle	$0.5 + 0.45 \cdot \sin\left(4\pi t/T\right)$
IV	Markov-Switching (MS)	$0.95 \cdot \mathbb{1}(s_{t-1} = 1) + 0.10 \cdot \mathbb{1}(s_{t-1} = 2)$
V	Logit Auto-Regressive (LAR)	$\alpha_t = 1/(1 + \exp\{-0.5x_t\})$
		$x_t = 0.97x_{t-1} + u_t, \qquad u_t \overset{iid}{\sim} \mathcal{N}(0, 1)$

4.2 Simulation Results

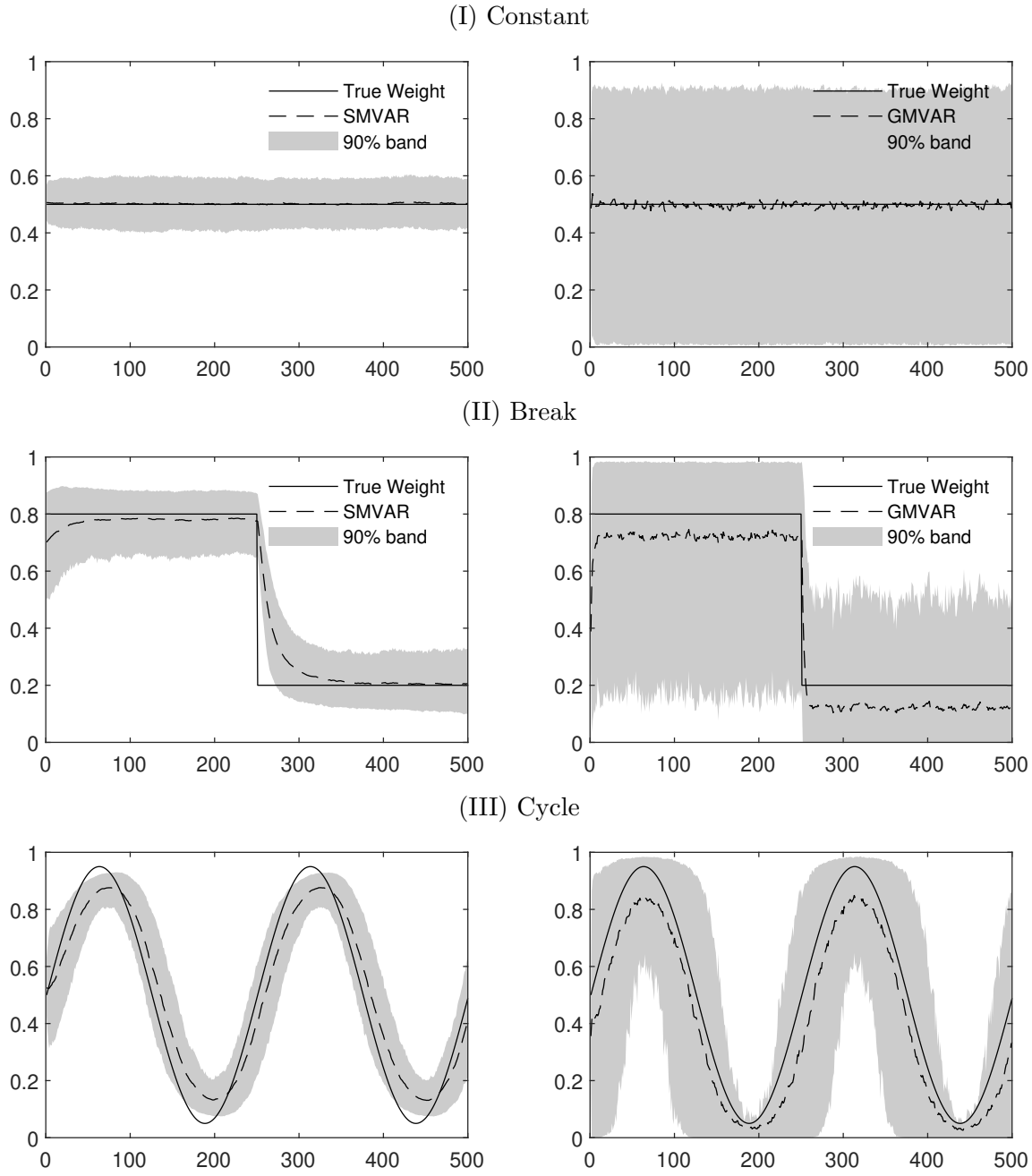
First, we discuss the performance of the SMVAR and the MVAR, MSVAR and GMVAR benchmarks with respect to filtering the mixture dynamics of DGPs I to V. Afterwards, we compare SMVAR and LMVAR in the presence of external factor-driven mixture dynamics.

4.2.1 Comparison of SMVAR and GMVAR

Figure 1 shows the average SMVAR and GMVAR estimates of α_t fitted to 1000 Monte Carlo replications alongside the true weights for Models I, II, and III.[8] The shaded areas indicate 90 percent confidence bands. The left panel of Figure 1 displays the results for the SMVAR, which shows similar adequate performance for all three DGPs. As all observation-driven models, the SMVAR needs some time to account for changes in the modeled process. Apart from that, we clearly see that the SMVAR is able to track the true mixture weight processes well. In particular, it can quite nicely adopt to structural breaks after a reasonable amount of periods.

[8]Note that a similar graphical representation of average estimates is not meaningful for the stochastic DGPs (Models IV and V) as the mixture weight paths change with every draw. We refer to Table 2 for an evaluation of these.

Figure 1: Filtered Mixture Weights



Notes: Figure shows filtered mixture weights $\alpha_{1,t}$ averaged over 1000 replication of the DGPs (I) Constant, (II) Break, and (III) Cycle. Solid lines display the true $\alpha_{1,t}$. The left (right) panel shows the averages of SMVAR (GMVAR) estimates as dashed lines. Shaded areas indicate 90 percent confidence bands.

The right panel of Figure 1 shows the result for the GMVAR. We document that the constant DGP I can be recovered on average, but the variation is particularly large when

compared to the SMVAR. The reason for this result can be deduced from the used updating mechanism in Eq. (12). To accommodate a constant $\alpha_{1,t}$, we would need to have $\alpha_1 = 0.5$ when abstracting from the unlikely and practically irrelevant scenario in which the ratio $p_1(y_{t-1})/p(y_{t-1})$ is constant over time. Hence, the DGP I with $\alpha_{1,t} \equiv \alpha_k = 0.5$ is not nested by the GMVAR. The SMVAR benefits here from the additional constant parameter ω that allows for incorporating the static case. The same problem is visible in the break process (II) that is estimated with huge confidence bands. The two constant levels cannot be captured correctly — even on average — because of the only parameter α_1 that has to serve as average weight and changing probability for drawing from state 1 at the same time. Interestingly, the level of the mixture weight is underestimated with almost the same size before and after the break. Hence, a constant in the mixture process, as included in the SMVAR, could be expected to crucially improve the GMVAR performance. An observation in favor of the GMVAR is that the break in DGP II is acknowledged much faster than by the SMVAR. The cycles are captured quite well by the GMVAR, although the confidence bands are still larger than for the SMVAR. In particular, the turning points are acknowledged faster. A reason for this faster detection of breaks and turning points is that the GMVAR depends only on the most recent observations and not on its own past value. This allows for an early detection of such structural changes, but comes with the drawback of lacking flexibility to capture persistent processes adequately.

Table 2 shows the average mean squared error (MSE) and the average mean absolute error (MAE) for the estimate of the mixture weight $\alpha_{1,t}$ from an MVAR, an MSVAR, a GMVAR, and an SMVAR model with 1000 Monte Carlo replications and different sample sizes T and Table B.1 in Appendix B the corresponding average goodness-of-fit statistics. We compute the error measures without the first 50 periods in order to mitigate the influence of starting values. The results are in line with the graphical inspection of Figure 1. The SMVAR shows a better performance than the GMVAR for all DGPs and sample lengths T. The improved performance is particularly visible for the constant mixture weight DGP I. It

Table 2: Average Estimation Error Comparison

		MSE				MAE			
DGP	T	MVAR	MSVAR	GMVAR	SMVAR	MVAR	MSVAR	GMVAR	SMVAR
	200	0.0064	0.0096	0.1276	0.0096	0.0619	0.0763	0.3205	0.0757
I Constant	500	0.0023	0.0032	0.1083	0.0041	0.0370	0.0448	0.2911	0.0511
	1000	0.0010	0.0015	0.0956	0.0027	0.0250	0.0305	0.2711	0.0430
	200	0.1077	0.0804	0.0881	0.0268	0.3005	0.2304	0.2313	0.1169
II Break	500	0.1029	0.0720	0.0797	0.0175	0.3000	0.2212	0.2184	0.0954
	1000	0.1017	0.0699	0.0781	0.0116	0.3000	0.2183	0.2145	0.0756
	200	0.1220	0.0790	0.0897	0.0339	0.3015	0.2222	0.2247	0.1410
III Cycle	500	0.1163	0.0720	0.0825	0.0184	0.2972	0.2133	0.2142	0.1040
	1000	0.1155	0.0690	0.0801	0.0127	0.2966	0.2089	0.2104	0.0875
	200	0.1938	0.0665	0.1039	0.0702	0.4093	0.1205	0.1755	0.1569
IV MS	500	0.1988	0.0664	0.1038	0.0669	0.4310	0.1156	0.1745	0.1387
	1000	0.2041	0.0665	0.1037	0.0663	0.4417	0.1134	0.1743	0.1317
	200	0.0865	0.0650	0.1030	0.0409	0.2442	0.2019	0.2434	0.1587
V LAR	500	0.0979	0.0664	0.0953	0.0385	0.2670	0.2076	0.2321	0.1549
	1000	0.1035	0.0677	0.0931	0.0383	0.2766	0.2107	0.2291	0.1547

Notes: Table shows the average mean squared error (MSE) and the average mean absolute error (MAE) for the estimate of the mixture weight $\alpha_{1,t}$ from a MVAR, a MSVAR, a GMVAR, and a SMVAR model with 1000 Monte Carlo replications and different sample sizes T.

is also worth noting that the performance of the SMVAR is considerably close to the one of the correctly specified MVAR benchmark. With respect to DGP IV, we see that the SMVAR is also able to track Markov-Switching mixture dynamics quite well with the errors being larger than the ones from using the correctly specified MSVAR, but considerably lower than the ones of the GMVAR and the static MVAR. With respect to DGP V, where the mixture weight is driven by an external VAR process, we also find that the SMVAR can best capture the in-sample weight dynamics.

Table 3 shows the mean squared prediction error (MSPE) and the mean absolute prediction error (MAPE) for the out-of-sample prediction of the mixture weight $\alpha_{1,T+1}$ from a MVAR, a MSVAR, a GMVAR, and a SMVAR model with 1000 Monte Carlo replications

Table 3: Average Prediction Error Comparison

		MSPE				MAPE			
DGP	T	MVAR	MSVAR	GMVAR	SMVAR	MVAR	MSVAR	GMVAR	SMVAR
	200	0.0064	0.0097	0.1277	0.0196	0.0619	0.0762	0.3209	0.1085
I Constant	500	0.0023	0.0033	0.1104	0.0065	0.0370	0.0454	0.2927	0.0639
	1000	0.0010	0.0015	0.0977	0.0033	0.0250	0.0305	0.2756	0.0455
	200	0.0634	0.0684	0.0788	0.0212	0.2268	0.2092	0.2346	0.1121
II Break	500	0.0445	0.0600	0.0615	0.0119	0.2027	0.1967	0.2123	0.0908
	1000	0.0418	0.0487	0.0518	0.0086	0.2002	0.1722	0.2034	0.0731
	200	0.0207	0.0626	0.1640	0.0450	0.1213	0.2175	0.3776	0.1801
III Cycle	500	0.0150	0.0583	0.1456	0.0285	0.1080	0.2127	0.3494	0.1397
	1000	0.0142	0.0521	0.1398	0.0178	0.1129	0.2040	0.3407	0.1089
	200	0.1985	0.0349	0.0683	0.0435	0.4148	0.0806	0.1372	0.1218
IV MS	500	0.1997	0.0286	0.0531	0.0332	0.4320	0.0719	0.1250	0.0995
	1000	0.2060	0.0306	0.0646	0.0333	0.4440	0.0725	0.1347	0.0933
	200	0.1077	0.0789	0.0899	0.0414	0.2731	0.2206	0.2253	0.1589
V LAR	500	0.1062	0.0715	0.0882	0.0346	0.2787	0.2197	0.2207	0.1454
	1000	0.1109	0.0753	0.0930	0.0351	0.2887	0.2228	0.2288	0.1479

Notes: Table shows the mean squared prediction error (MSPE) and the mean absolute prediction error (MAPE) for the out-of-sample prediction of the mixture weight $\alpha_{1,T+1}$ from a MVAR, a MSVAR, a GMVAR, and a SMVAR model with 1000 Monte Carlo replications and different sample sizes T.

and different sample sizes T. The results show that the performance of the SMVAR in comparison to the GMVAR also holds out-of-sample.[9]

As bottom line of this Monte Carlo study, we can conclude that the SMVAR is able to recover a variety of different mixture dynamics and performs better than its dynamic competitor model, the GMVAR both, in-sample and out-of-sample. The latter has nice theoretical properties because of its particular updating scheme but lacks the flexibility to accommodate a wide range of practically relevant DGPs, which the SMVAR is able to handle adequately by adding a constant and an autoregressive component in the mixture weight updating scheme.

[9]The out-of-sample MVAR and MSVAR results for DGPs I and III should not be overinterpreted as the weighting process approaches its average value in T by construction.

4.2.2 Comparison of SMVAR and LMVAR

The LMVAR of Burgard et al. (2019) uses covariates to explain the mixture weight dynamics. This is particular useful if such covariates are known or proposed by theory. However, it is evident that this method is not favorable if such covariates are not available or only poor signals of the true drivers. In the following, we will examine the case in which the explanatory value of the external factor is a priori uncertain.

We simulate from the same VAR model from Eq. (19) with the mixture process V. Again, we use 1000 Monte Carlo replications of different sample sizes $T = 200, 500, 1000$. We then compare the estimates of a SMVAR model with those of a LMVAR. Note that the dynamics of α_t follow an LMVAR model in Eq. (13) with $K = 2$. Hence, if we estimate an LMVAR given the true covariate x_t, it will trivially outperform our model. The interesting question is what happens if we provide the LMVAR estimator with a noisy signal $\tilde{x}_t$ of x_t. We derive this signal with

$$\tilde{x}_t = \rho x_t + (1 - \rho)\tilde{u}_t, \qquad \tilde{u}_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right) \tag{20}$$

where $\sigma^2 = 16.9205$ is equal to the unconditional variance of the process x_t. Hence, the information content of the signal ranges from pure noise ($\rho = 0$) to perfect information ($\rho = 1$).

Table 4 shows in panels (a) and (b) the average MSE and the average MAE for the estimate of the mixture weight $\alpha_{1,t}$ from a SMVAR and an LMVAR model. We see clearly that the LMVAR using the true covariate outperforms all models (as expected). More revealing are the cases where the SMVAR outperforms the LMVAR, that is, when the signal is noisy ($\rho = 0, 0.2, 0.4$). The breakeven point for which both methods are on par can be located slightly above a ρ of 0.5. Hence, over 50% of the signal's variation must come from the true covariate to justify using the LMVAR in this setting. A similar result is found for the out-of-sample measures in panels (c) and (d) of table 4, where the breakeven value of noise

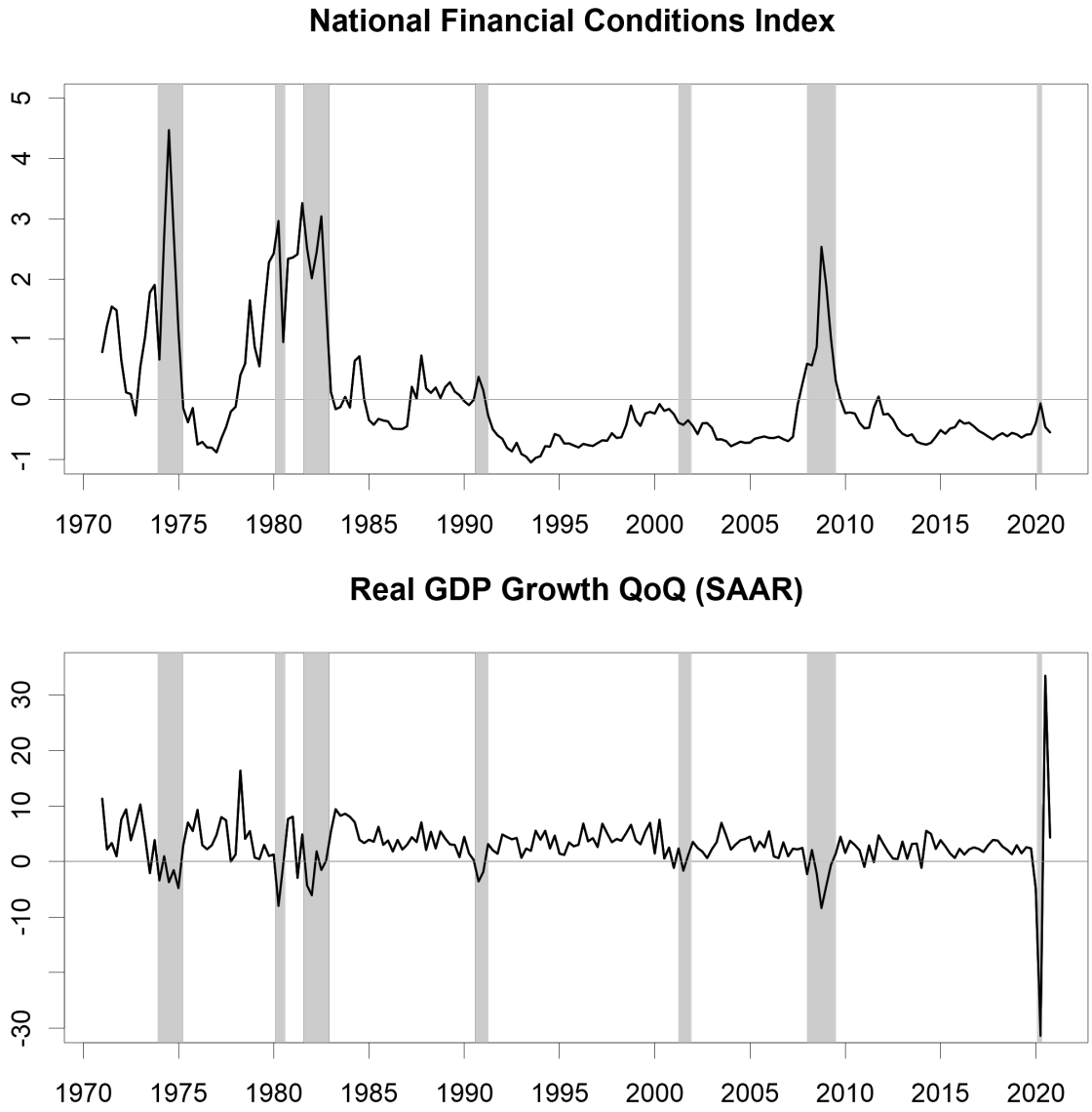
Table 4: Comparison of Average Estimation Errors: SMVAR vs. LMVAR

T	SMVAR	LMVAR					
		$\rho = 0$	0.2	0.4	0.6	0.8	1
(a) MSE							
200	0.0409	0.0893	0.0854	0.0685	0.0370	0.0126	0.0055
500	0.0385	0.0985	0.0933	0.0709	0.0349	0.0087	0.0016
1000	0.0383	0.1038	0.0979	0.0731	0.0346	0.0078	0.0008
(b) MAE							
200	0.1587	0.2481	0.2416	0.2115	0.1482	0.0827	0.0543
500	0.1549	0.2675	0.2591	0.2205	0.1453	0.0682	0.0298
1000	0.1547	0.2767	0.2675	0.2252	0.1447	0.0641	0.0212
(c) MSPE							
200	0.0414	0.1109	0.1074	0.0867	0.0434	0.0125	0.0053
500	0.0346	0.1074	0.1026	0.0795	0.0389	0.0088	0.0016
1000	0.0351	0.1113	0.1035	0.0750	0.0339	0.0072	0.0008
(d) MAPE							
200	0.1589	0.2769	0.2713	0.2378	0.1595	0.0808	0.0527
500	0.1454	0.2799	0.2727	0.2349	0.1542	0.0692	0.0301
1000	0.1479	0.2890	0.2770	0.2282	0.1425	0.0613	0.0209

Notes: Table shows the average mean squared error (MSE), the average mean absolute error (MAE), the mean squared prediction error (MSPE) and the average mean absolute prediction error (MAPE) for the estimate of the mixture weight $\alpha_{1,t}$ from a SMVAR and an LMVAR model with 1000 Monte Carlo replications and different sample sizes T. The parameter ρ indicates the informational content of the covariate provided to the LMVAR.

improves slightly in favor of the SMVAR. Noise ratios of 50% or more are not uncommon for many economic and, in particular, financial variables. Hence, the main takeaway from this exercise is that the SMVAR is not only favorable if no adequate covariates are known. In addition, it can serve as a worthwhile complement to an LMVAR analysis if the considered covariates are of questionable quality.

Figure 2: NFCI and Real GDP Growth



Notes: Positive (negative) values of the normalized NFCI indicate tighter-than-average (looser-than-average) financial conditions. Real GDP Growth QoQ is shown in percentage. Shaded areas are NBER recessions.

5 Empirical Application

Our empirical application aims at highlighting real-financial linkages. Our variable selection is inspired by Adrian et al. (2021) who document multimodality in the conditional distribution of economic growth, especially in times with tight financial conditions. Hence, we choose the quarter-over-quarter annualized growth rate of real GDP as indicator for real economic

activity.[10] To capture financial conditions, we rely on the National Financial Conditions Index (NFCI) provided by the Federal Reserve Bank of Chicago.[11] The NFCI is a weighted average of 105 indicators of risk, credit, and leverage in the financial system, each expressed relative to its sample average and scaled by its sample standard deviation (SD). Positive (negative) values of the NFCI are associated with tighter-than-average (looser-than-average) financial conditions. Our sample starts in 1971q1 with the first observation of the NFCI and ends in 2020q4.[12]

Figure 2 shows both series over time. All (gray-shaded) NBER recessions are accompanied by positive values of the NFCI with the mild and short recession of 2001 being an exception. In addition, the Covid-19 recession is also only associated with a short-lived surge in the NFCI. Real GDP growth is lower and financial conditions tighten during recessions, so that economic and financial conditions are negatively correlated ($\rho = -0.24$; $\rho = -0.33$ when excluding the year 2020).

5.1 Empirical Results

We estimate a two-state SMVAR with $p = q = 1$ to examine the joint distribution of the NFCI and real GDP. Both AIC and BIC are in favor of one lag in each of the two regimes. We estimate a slightly re-parameterized internal updating process for the unrestricted mixture weights that is given by

$$\tilde{\alpha}_{t+1} = \overline{\tilde{\alpha}} + a\frac{\exp(\tilde{\alpha}_t)}{(1+\exp(\tilde{\alpha}_t))^2}\frac{p_1(y_t|\mathcal{F}_{t-1}) - p_2(y_t|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})} + b\left(\tilde{\alpha}_t - \overline{\tilde{\alpha}}\right). \tag{21}$$

where $\overline{\tilde{\alpha}} = \omega/(1-b)$ is the unconditional mixture weight in case the process is stationary.

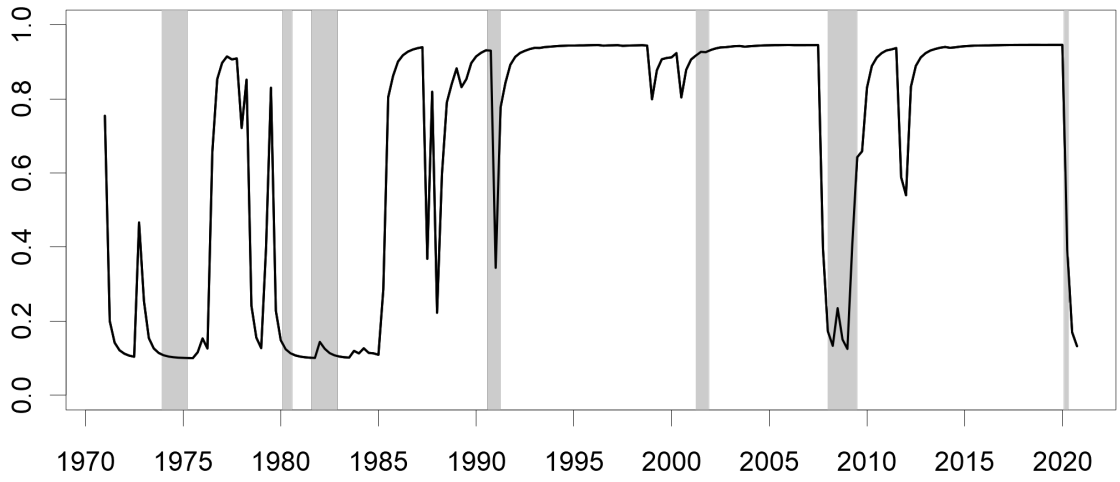
Figure 3 shows the evolution of the transformed mixture weights $\alpha_t = \exp(\tilde{\alpha}_t)/[1 + \exp(\tilde{\alpha}_t)]$ over time. All recessions (except the one in 2001) are characterized by a low value

[10] Source: https://fred.stlouisfed.org/series/GDPC1.

[11] The data set and some background can be found here: https://www.chicagofed.org/publications/nfci/index.

[12] The null hypothesis of non-stationarity can be rejected at the 1% level for both series.

Figure 3: Development of α_t over Time



Notes: Figure shows the probability of being in the normal state as indicated by the estimated mixture weights α_t. Shaded areas are NBER recessions.

of the mixtures weights α_t. In addition, the recessions are captured in a timely manner and even anticipated in the case of both oil crises in the 1970s and early-1980s, the Global Financial Crisis of 2008−2009, and the Covid-19 slump of 2020. Only the recession of 1990−1991 is captured with a minimal delay of two quarters. Since the latter recession and the one of 2001 are accompanied with only a minor tightening of financial conditions (if at all), we can interpret the mixture weights as indicator for *joint* economic and financial conditions. This is reassuring as the aim of the SMVAR is to capture the joint distribution of the two variables.

Table 5 shows the coefficients of the SMVAR model and the (co-)variances for both states as well as the estimates for GAS mixture weight updating procedure of Eq. (10). Standard errors can be found in parentheses. The variance parameters $\Omega_{1,1}$ for the NFCI and $\Omega_{2,2}$ for real GDP growth indicate that the regime on the right-hand side of Table 5 is the more volatile state. In addition, we observe a positive value for the constant of the NFCI equation in this state, whereas the corresponding value for the other state is negative. Put differently, the financial conditions are, ceteris paribus, tighter in the regime on the right-hand side. Hence, we can interpret this as crisis regime. The NFCI is similarly persistent in both

Table 5: Estimation Results

(a) VAR Parameters

	Normal Regime		Crisis Regime	
	$NFCI_t$	ΔGDP_t	$NFCI_t$	ΔGDP_t
$NFCI_{t-1}$	0.7490	0.0162	0.7506	0.8000
	(0.0268)	(0.3418)	(0.0858)	(1.5886)
ΔGDP_{t-1}	−0.0134	0.1058	−0.0211	−0.2059
	(0.0033)	(0.0524)	(0.0082)	(0.1485)
const	−0.1433	0.0203	0.3026	0.0507
	(0.0163)	(0.0023)	(0.1366)	(0.0130)

(b) (Co-)Variance Parameters

	Normal Regime	Crisis Regime
$\Omega_{1,1}$	0.1168	0.7647
	(0.0085)	(0.0704)
$\Omega_{1,2}$	0.0025	-0.0089
	(0.0015)	(0.0085)
$\Omega_{2,2}$	0.0177	0.0723
	(0.0011)	(0.0067)

(c) GAS Parameters

$\overline{\alpha}$	1.1230
	(0.8108)
a	3.3267
	(0.7807)
b	0.9006
	(0.0459)

Notes: Table shows the coefficients of the SMVAR model and the (co-)variances for both states as well as the estimates for GAS mixture weight updating procedure of Eq. (10). Standard errors are in parentheses.

states. Changes in last period's real GDP growth lead to a decrease in the current period's NFCI with a numerically larger effect in the crisis regime. Finally, GDP is found to be

persistent in the normal state, whereas the effect is negative (albeit insignificant) during crisis times.[13]

Turning to the GAS parameters, we find a high degree of persistence in the weights as indicated by the estimate of 0.90 for the AR part b of Eq. (21). This is also the reason for the smooth development of α_t over time. The coefficient for the weight update part a is positive and significant, implying that the scaled observation density evaluated at the current observation is informative for detecting changes in future mixture weights. Finally, the estimate for the average mixture weight $\overline{\alpha} = \exp\left(\overline{\tilde{\alpha}}\right) / \left(1 + \exp\left(\overline{\tilde{\alpha}}\right)\right)$ is 0.75. Consequently, the economy is, on average, in 75% of the time in the normal state.

5.2 Impulse Response Analysis

We further elaborate on macro-financial linkages with a generalized impulse response analysis. Shocks are recursively identified with a Cholesky decomposition where we rule out an instantaneous impact of financial shocks on the relatively slow moving GDP growth series. We consider two types of adverse shocks: (i) an economic shock that decreases GDP growth by four SD and (ii) a financial shock that increases the NFCI by four SD. GIRFs are simulated according to Algorithm 1 with $R = 10000$ Monte Carlo replications. The pre-shock observation of the time series is assumed to be equal to the unconditional mean in the assumed initial regime.

In a first exercise, we look at the component-specific GIRFs for which we rule out regime shifts. The black lines in the left panel of Figure 4 show the mean responses in the normal times regime under the assumption that no shift into the crisis regime may occur. The grey-shaded areas indicate 68% confidence bands. We find that the impact of the economic shock on GDP growth fades out quickly, which is no surprise since we use quarter-over-quarter growth rates. The adverse economic shock significantly reduces the NFCI on impact with a peak effect of -0.076 SD. With respect to the financial shock, we see a greater degree of

[13] This negative coefficient might also be the reason of why the GDP intercept is larger in the crisis state.

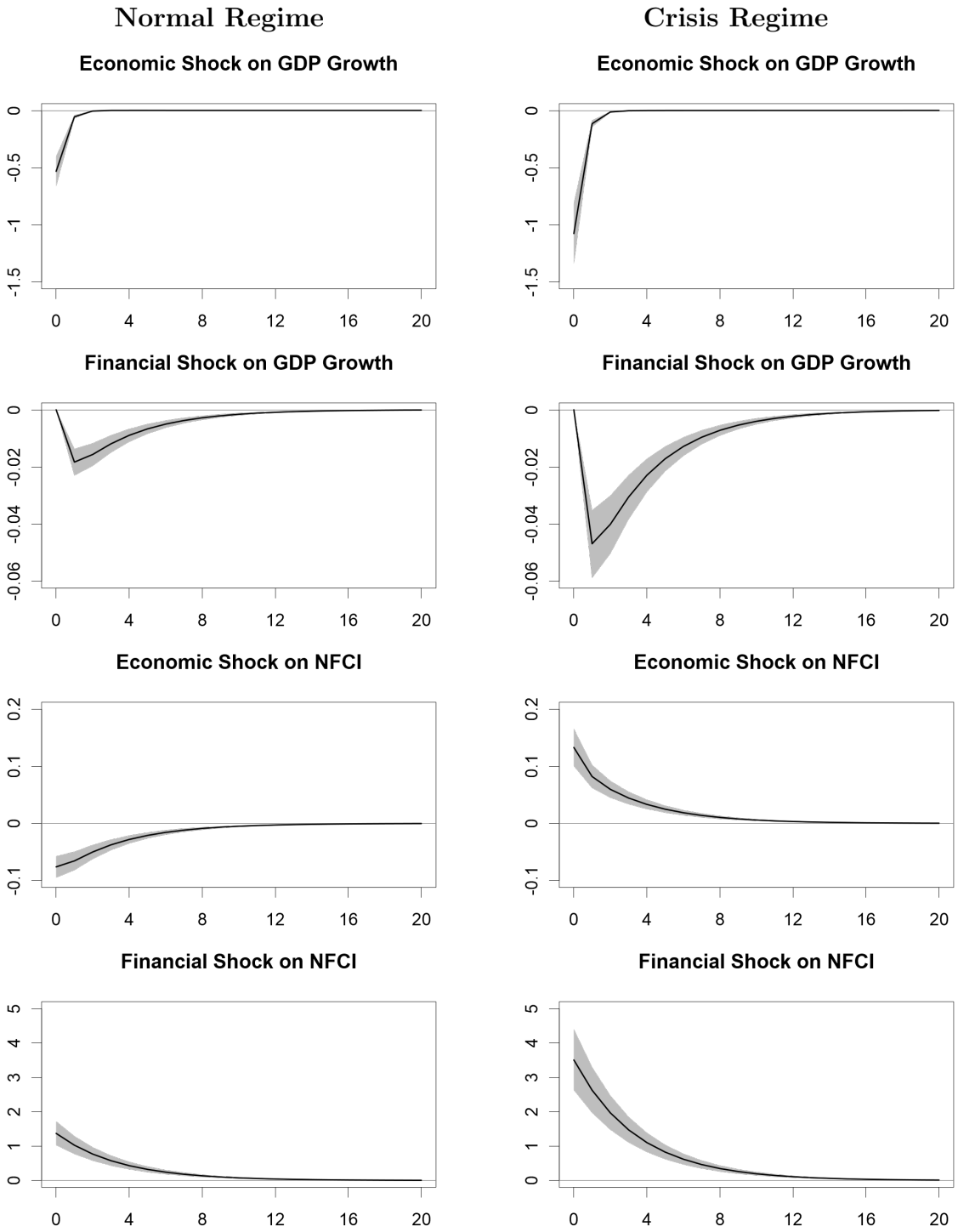
persistence as its impact dies out only slowly. The financial shock also reduces GDP growth significantly with a peak effect of -1.8 basis points (bps).

The right panel of Figure 4 shows the GIRFs that occur within the crisis regime. Again, we rule out a shift into the other regime. The qualitative behavior of the impulse responses is very similar to the behavior in the normal state. In terms of magnitude, we find a significantly larger peak response of real GDP growth (-4.7 bps as compared to -1.8 bps in the normal regime) after a financial shock. One reason for this difference might be the size of the financial shock, which is roughly 2.5 times larger in the crisis regime. However, there is one crucial difference that cannot be explained by the relative shock size across regimes. The impulse response of the NFCI to the economic shock has the opposite sign. It appears that while an adverse economic shock leads to a slight easing of financial conditions in normal times, the opposite can be found in a crisis. Here, we find a peak effect of 0.133 SD (as opposed to -0.076 SD in the normal regime). One reason for this could be that countermeasures against economic shocks also improve financial conditions in normal times, whereas this is not the case in crisis times. This holds in particular if the economic shock does not trigger a change into the crisis regime (as ruled out by assumption).

Looking at the component-wise impulse responses — which are equivalent to those of linear VAR models — is one way to analyze time series dynamics. However, the SMVAR explicitly models regime probabilities that recursively depend on prior observations. Hence, the SMVAR not only allows to take the possibility of regime shifts into account, but also to analyze the impact of shocks on the mixture weights. Consequently, we can investigate the impact of an economic or financial shock on the probability of a switch into the crisis regime. Figure 5 shows the GIRFs to an economic and a financial shock of four SD starting in the normal regime.

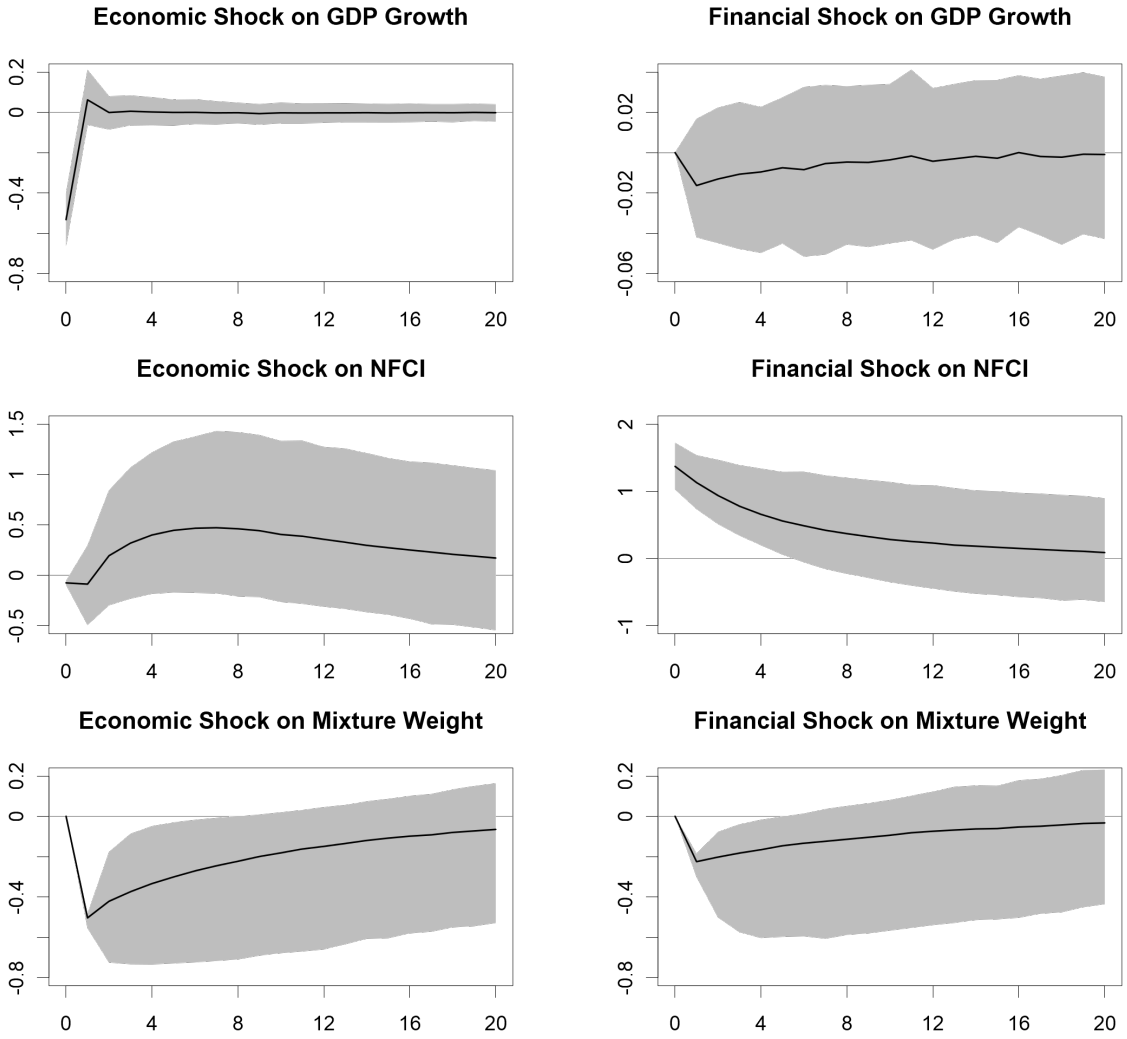
There are four striking findings worth discussing. First, the qualitative results concerning the persistence of the shocks are similar when introducing possible regime shifts. The economic (financial) shock decreases (increases) the shock variable significantly on impact

Figure 4: Component-wise Impulse Responses



Notes: Figure shows generalized impulse responses (black lines) to an economic and a financial shock of four SD within the normal regime (left panel) and the crisis regime (right panel). Regime shifts are ruled out. Grey-shaded areas indicate 68% confidence bands.

Figure 5: Impulse Responses with Mixture Dynamics



Notes: Figure shows generalized impulse responses (black lines) to an economic and a financial shock of four SD starting in the normal regime. Grey-shaded areas indicate 68% confidence bands.

and the effect is rapidly (slowly) dying out. Similarly, GDP growth decreases shortly after a financial shock with a similar pattern as in the case of both component-wise impulse responses.

Second, the confidence bands largely increase for all impulse responses. This phenomenon can be observed for dynamic MVAR models in general. The impulse responses largely differ depending on whether the initial shock triggers a regime shift or not. In our case, we see in the lower panel of Figure 5 that both shocks significantly decrease the mixture weight and

therefore make a shift into the crisis regime more likely. Subsequently, the trajectories of the variables may strongly differ depending on whether a crisis was triggered or not. Confidence bands may therefore be much less informative as compared to a linear VAR but still provide some insight into the range of possible variable developments.

Third, we also notice that the likelihood of triggering a crisis is, on average, higher after an economic shock than after a financial shock of equal relative size. More precisely, the probability of a crisis increases by roughly 50 percentage points (pp) one period after a shock to economic growth as compared to 22 pp in the case of a financial shock. Given an initial normal times regime weight of 90%, the probability of switching into the crisis increases from 10% to 60% (32 %) in the case of an economic (financial) shock.

Fourth, the response of the NFCI after the economic shock largely differs from both component-wise responses that have opposing signs. The response in the full mixture model is negative on impact in a magnitude similar to the normal regime but then turns positive with a peak effect that is ten times larger than the peak effect of the same shock in the crisis regime. Moreover, the shock lasts much longer in the full mixture model. This observed dynamic is an example of how nonlinear VAR models — like the SMVAR — are able to alter our understanding of responses to shocks. In the regime-specific impulse responses, the linkage between both series is underestimated so that the growth shock has only a small effect on the NFCI that vanishes quickly. However, in the SMVAR the initial growth shock not only transmits directly though the normal times VAR model connections, but additionally increases the probability of shifting into a crisis. If the economy switches into the crisis regime (with a probability of around 60%), the subsequent trajectory will be affected by the higher unconditional NFCI mean and the higher variance within the crisis regime. Conversely, with a probability of 40% the economy will remain in the normal regime having a trajectory in which the shock dies out after a few quarters. Taken together, this explains the high average impact on the NFCI and the large confidence bands.

6 Conclusions

We proposed a novel dynamic mixture vector autoregressive model with time-varying mixture weights, which are driven by earlier observations of the endogenous variables, the Score-Driven Mixture Vector Autoregression. Our weight updating scheme follows the generalized autoregressive score (or score-driven) approach developed by Creal et al. (2013) and Harvey (2013). The derived scheme uses the scaled conditional density of the VAR component model, evaluated at the current observation. Intuitively, the procedure increases the state weight for the k-th component of the VAR model in the following period if the current observation is more likely to be drawn from this particular state. The SMVAR is more flexible than other dynamic MVAR approaches and allows for straightforward likelihood-based estimation and inference.

We performed a Monte Carlo study to investigate the ability of the SMVAR to recover mixture dynamics from various data generating processes and find that the SMVAR outperforms the GMVAR in filtering mixture weights with non-Markovian dynamics both, in-sample and out-of-sample. One reason for this is the more flexible specification of our updating scheme as the SMVAR includes a constant and an autoregressive part in the mixture weight dynamics, whereas the Gaussian MVAR always updates the weights according to the recent scaled observation density. In addition, the SMVAR is also helpful if the employed signal driving the mixture weights in a Logit MVAR is uncertain. It can serve as a hedge against potential misspecification and outperforms the Logit MVAR if the signal is too noisy.

In an empirical application, we modeled the joint distribution of the National Financial Conditions Index and real GDP growth using a two-state SMVAR. We showed that the mixture weights disentangle a (tranquil) normal regime and a (volatile) economic and financial crisis regime. In addition, almost all NBER recessions are accompanied by large values for the crisis state and also anticipated by a drastic change in the mixture weights. Our application highlights once more the appealing feature of utilizing an AR term in the mixture weight dynamics, which leads to a smooth development of the weights over time. General-

ized impulse response functions based on Koop et al. (1996) revealed that adverse shocks on economic growth and financial conditions significantly increase the conditional probability for switching into a crisis regime for several quarters with economic shocks having a larger impact. The increased switching probability enlarges the set of possible trajectories after a shock and considerably alters the impact of the growth shock on financial conditions as compared to the component-wise impulse response where a regime switch has been ruled out.

References

Adrian, T., Boyarchenko, N., Giannone, D., 2021. Multimodality in macrofinancial dynamics. International Economic Review 62, 861–886.

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, 458–478.

Bennani, H., Burgard, J.P., Neuenkirch, M., 2022. The Financial Accelerator in the Euro Area: New Evidence Using a Mixture VAR Model. Macroeconomic Dynamics (forthcoming).

Bernardi, M., Catania, L., 2019. Switching generalized autoregressive score copula models with application to systemic risk. Journal of Applied Econometrics 34, 43–65.

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

Burgard, J.P., Neuenkirch, M., Nöckel, M., 2019. State-Dependent Transmission of Monetary Policy in the Euro Area. Journal of Money, Credit and Banking 51, 2053–2070.

Camacho, M., 2004. Vector smooth transition regression models for US GDP and the composite index of leading indicators. Journal of Forecasting 23, 173–196.

Catania, L., 2021. Dynamic adaptive mixture models with an application to volatility and risk. Journal of Financial Econometrics 19, 531–564.

Creal, D., Koopman, S.J., Lucas, A., 2013. Generalized Autoregressive Score Models with Applications. Journal of Applied Econometrics 28, 777–795.

Fong, P.W., Li, W.K., Yau, C.W., Wong, C.S., 2007. On a Mixture Vector Autoregressive Model. Canadian Journal of Statistics 35, 135–150.

Gorgi, P., Hansen, P.R., Janus, P., Koopman, S.J., 2019. Realized Wishart-GARCH: A Score-driven Multi-Asset Volatility Model. Journal of Financial Econometrics 17, 1–32.

Hamilton, J., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357–84.

Hamilton, J., 1990. Analysis of time series subject to changes in regime. Journal of Econometrics 45, 39–70.

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

Harvey, A.C., Lange, R.J., 2017. Volatility modeling with a generalized t distribution. Journal of Time Series Analysis 38, 175–190.

Kalliovirta, L., Meitz, M., Saikkonen, P., 2016. Gaussian mixture vector autoregression. Journal of Econometrics 192, 485–498.

Koop, G., Pesaran, M.H., Potter, S.M., 1996. Impulse response analysis in nonlinear multivariate models. Journal of Econometrics 74, 119–147.

Krolzig, H.M., 1997. Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin Heidelberg.

Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Annals of Mathematical Statistics 22, 79–86.

Oh, D.H., Patton, A.J., 2018. Time-Varying Systemic Risk: Evidence From a Dynamic Copula Model of CDS Spreads. Journal of Business & Economic Statistics 36, 181–195.

Tsay, R.S., 1998. Testing and modeling multivariate threshold models. Journal of the American Statistical Association 93, 1188–1202.

Ugray, Z., Lasdon, L., Plummer, J., Glover, F., Kelly, J., Martí, R., 2007. Scatter search and local NLP solvers: A multistart framework for global optimization. INFORMS Journal on Computing 19, 328–340.

Virolainen, S., 2020. Structural Gaussian mixture vector autoregressive model. arXiv preprint arXiv:2007.04713 .

Weise, C.L., 1999. The asymmetric effects of monetary policy: A nonlinear vector autoregression approach. Journal of Money, Credit and Banking 31, 85–108.

Appendix

A Derivations for GAS Updating Equations

A.1 General weight updating

$$\nabla_t = \frac{\partial}{\partial \tilde{\alpha}_t} \ln p(y_t|\mathcal{F}_{t-1}) \tag{A.1}$$

$$= \frac{\partial}{\partial \tilde{\alpha}_t} \ln \sum_{k=1}^{K} \alpha_{k,t} p_k(y_t|\mathcal{F}_{t-1}) \tag{A.2}$$

$$= \frac{\partial h}{\partial \tilde{\alpha}_t} \cdot \frac{\partial}{\partial \alpha_t} \ln \sum_{k=1}^{K} \alpha_{k,t} p_k(y_t|\mathcal{F}_{t-1}) \tag{A.3}$$

$$= \mathcal{J}_h(\tilde{\alpha}_t) \cdot \frac{\frac{\partial}{\partial \alpha_t} \sum_{k=1}^{K} \alpha_{k,t} p_k(y_t|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})} \tag{A.4}$$

$$= \mathcal{J}_h(\tilde{\alpha}_t) \cdot \begin{pmatrix} \frac{p_1(y_t|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})} \\ \vdots \\ \frac{p_K(y_t|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})} \end{pmatrix} \tag{A.5}$$

A.2 Jacobian of logit transformation

$\underline{k = l, k \neq K}$

$$\frac{\partial h_k}{\partial \tilde{\alpha}_l}(\tilde{\alpha}) = \frac{\partial}{\partial \tilde{\alpha}_l} \frac{\exp(\tilde{\alpha}_k)}{1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)} \tag{A.6}$$

$$= \frac{\exp(\tilde{\alpha}_k)\left(1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)\right) - \exp(\tilde{\alpha}_k)\exp(\tilde{\alpha}_l)}{\left(1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)\right)^2} \tag{A.7}$$

$$= \frac{\exp(\tilde{\alpha}_k)\left(1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)\right) - \exp(2\tilde{\alpha}_k)}{\left(1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)\right)^2} \tag{A.8}$$

$\underline{k \neq l, k \neq K}$

$$\frac{\partial h_k}{\partial \tilde{\alpha}_l}(\tilde{\alpha}) = \frac{\partial}{\partial \tilde{\alpha}_l} \frac{\exp(\tilde{\alpha}_k)}{1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)} \tag{A.9}$$

$$= \frac{-\exp(\tilde{\alpha}_k)\exp(\tilde{\alpha}_l)}{\left(1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)\right)^2} \tag{A.10}$$

$$\tag{A.11}$$

$\underline{k = K}$

$$\frac{\partial h_k}{\partial \tilde{\alpha}_l}(\tilde{\alpha}) = \frac{\partial}{\partial \tilde{\alpha}_l} \left(1 - \sum_{j=1}^{K-1} \frac{\exp(\tilde{\alpha}_j)}{1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)}\right) \tag{A.12}$$

$$= -\frac{\partial}{\partial \tilde{\alpha}_l} \frac{\exp(\tilde{\alpha}_l)}{1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)} - \sum_{j=1, j\neq l}^{K-1} \frac{\partial}{\partial \tilde{\alpha}_l} \frac{\exp(\tilde{\alpha}_j)}{1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)} \tag{A.13}$$

$$= \frac{-\exp(\tilde{\alpha}_l)\left(1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)\right) + \exp(2\tilde{\alpha}_l) + \sum_{j=1, j\neq l}^{K-1} \exp(\tilde{\alpha}_j)\exp(\tilde{\alpha}_l)}{\left(1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)\right)^2} \tag{A.14}$$

$$= \frac{-\exp(\tilde{\alpha}_l)}{\left(1 + \sum_{i=1}^{K-1} \exp(\tilde{\alpha}_i)\right)^2} \tag{A.15}$$

B Additional Simulation Results

Table B.1: Average Goodness-of-Fit Statistics

		Log Likelihood				AIC				BIC			
DGP	T	MVAR	MSVAR	GMVAR	SMVAR	MVAR	MSVAR	GMVAR	SMVAR	MVAR	MSVAR	GMVAR	SMVAR
	200	-430.93	-428.23	-450.15	-430.19	899.85	898.46	938.29	902.38	962.52	967.72	1000.96	971.64
I Constant	500	-1093.16	-1090.55	-1143.95	-1093.14	2224.32	2223.10	2325.89	2228.27	2304.40	2311.61	2405.97	2316.78
	1000	-2194.15	-2191.44	-2297.94	-2195.31	4426.31	4424.88	4633.88	4432.62	4519.56	4527.94	4727.13	4535.68
	200	-433.23	-422.77	-435.51	-421.51	904.46	887.54	909.02	885.02	967.13	956.80	971.69	954.29
II Break	500	-1091.66	-1077.14	-1101.81	-1058.89	2221.32	2196.29	2241.62	2159.78	2301.40	2284.79	2321.70	2248.29
	1000	-2188.00	-2167.28	-2211.06	-2110.15	4414.00	4376.56	4460.12	4262.31	4507.24	4479.62	4553.37	4365.37
	200	-433.62	-419.86	-432.89	-418.79	905.23	881.72	903.78	879.58	967.90	950.98	966.44	948.84
III Cycle	500	-1090.38	-1070.02	-1092.51	-1045.51	2218.76	2182.03	2223.02	2133.02	2298.84	2270.54	2303.10	2221.52
	1000	-2184.04	-2153.31	-2191.33	-2088.58	4406.07	4348.62	4420.65	4219.17	4499.32	4451.68	4513.90	4322.23
	200	-373.42	-346.28	-362.89	-351.66	784.84	734.56	763.77	745.33	847.51	803.83	826.44	814.59
IV MS	500	-960.97	-889.52	-925.77	-896.47	1959.93	1821.04	1889.54	1834.94	2040.01	1909.54	1969.62	1923.44
	1000	-1931.61	-1789.01	-1857.06	-1798.59	3901.22	3620.01	3752.12	3639.19	3994.47	3723.08	3845.36	3742.25
	200	-421.86	-415.52	-430.93	-414.32	881.72	873.05	899.87	870.65	944.39	942.31	962.53	939.91
V LAR	500	-1083.25	-1070.73	-1100.53	-1057.16	2204.50	2183.47	2239.06	2156.31	2284.57	2271.97	2319.14	2244.82
	1000	-2182.35	-2158.59	-2211.96	-2123.11	4402.69	4359.17	4461.93	4288.22	4495.94	4462.23	4555.17	4391.28

Notes: Table shows the average log likelihood, Akaike Information Criterion and the Bayesian Information Criterion for the estimate of the mixture weight $\alpha_{1,t}$ from a MVAR, a MSVAR, a GMVAR, and a SMVAR model with 1000 Monte Carlo replications and different sample sizes T.