

**Alexander Georges Gretener**  
**Matthias Neuenkirch**  
**Dennis Umlandt**

**Dynamic Mixture Vector Autoregressions  
with Score-Driven Weights**

**Research Papers in Economics**  
**No. 2/22**

# Dynamic Mixture Vector Autoregressions with Score-Driven Weights\*

Alexander Georges Gretener

University of Kiel

Matthias Neuenkirch

Trier University and CESifo

Dennis Umlandt

University of Innsbruck

November 27, 2024

## Abstract

We propose a novel dynamic mixture vector autoregressive (VAR) model in which the time-varying mixture weights are driven by the predictive likelihood score. Intuitively, the state weight of the  $k$ -th component VAR model is increased in the subsequent period if the current observation is more likely to have been drawn from this particular state. The model is not limited to a specific distributional assumption and allows for straightforward 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 handle appropriately. Finally, the empirical performance of the approach is illustrated by two applications: (i) the conditional joint distribution of stock and bond returns, and (ii) the regime-dependent connection of economic and financial conditions.

**Keywords:** Dynamic Mixture Models; Generalized Autoregressive Score Models; Macro-Financial Linkages; Nonlinear Vector Autoregressions; Stock and Bond Return Dynamics.

**JEL Codes:** C32; C34; G17.

---

\*The authors gratefully acknowledge comments and suggestions by two anonymous reviewers, Jan Pablo Burgard, Markus Haas, and the 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 commonly used 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 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 a linear Gaussian VAR process weighted by 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 governed by 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 **Score-driven Mixture Vector Autoregression** (SMVAR), we follow the generalized autoregressive score (GAS)<sup>1</sup> 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.<sup>2</sup> 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

---

<sup>1</sup>Also referred to as score-driven (SD) model or dynamic conditional score (DCS) model.

<sup>2</sup>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. Since the updating sequence only depends on the observation densities, the framework can conveniently accommodate specifications in which some of the components are VAR models while others are not.<sup>3</sup> Interestingly, this updating mechanism is similar to those of 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 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 specified freely.

In addition, we introduce the SMXVAR model which extends the weight updating process of the baseline model by additionally incorporating external predictors. We discuss the computation of generalized impulse responses of the model variables as well as the mixture weights. The (conditional) likelihood function of the SM(X)VAR model can be evaluated directly and allows for straightforward likelihood-based estimation and inference.

First, we perform a Monte Carlo study to investigate the abilities and limitations of several dynamic MVAR models to recover mixture dynamics from various data-generating processes both in-sample and out-of-sample. We find that the SM(X)VAR outperforms the benchmark models across several deterministic and stochastic processes.

Second, we illustrate the practicality of the flexible SM(X)VAR model using two empirical applications with time series of different frequency and noisiness. In Application 1, we investigate the relation between monthly stock and bond returns, following previous studies in the context of Markov-switching models (Guidolin and Ono, 2006; Guidolin and Timmermann, 2006; Kole and van Dijk, 2023). The resulting estimated SMVAR features two distinct components with differing levels of volatility. In line with prior contributions, stock

---

<sup>3</sup>In this regard, the SMVAR can be interpreted as a special case of the score-driven dynamic mixture model as proposed in the online appendix of Creal et al. (2013).

and bond returns are more strongly correlated in the high volatility component. The weight changes in the mixture density are moving slowly and cyclically, indicating a time-varying mixture rather than a (strict) regime interpretation. The SMXVAR results indicate that in particular term structure variables can predict weight changes. However, the score-driven updating remains the most important driver of weight dynamics when including external predictors.

In Application 2, we model the joint distribution of the National Financial Conditions Index and real GDP growth 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. In an impulse response analysis, we find that adverse shocks on economic growth and financial conditions significantly increase the conditional probability of switching into the crisis regime for several quarters with financial shocks having a larger impact.

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 applications are presented in Section 5. Section 6 concludes.

## 2 (Dynamic) Mixture Vector Autoregressions

Let  $(y_t)_{t=1}^{\infty}$  denote a  $d$ -dimensional time series defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , equipped with a filtration  $\mathcal{F}_t = \sigma(\{y_t, \dots, y_1\})$  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  $s_t = (s_{1,t}, \dots, s_{K,t})^\top$  where in every period  $t$  either  $s_{k,t} = 1$  (state  $k$  is active in period  $t$ ) or  $s_{k,t} = 0$  holds ( $k = 1, \dots, K$  and  $\sum_{k=1}^K s_{k,t} = 1$ ). A general MVAR model with  $K$  mixture components is then given by

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

where  $(\varepsilon_t)_{t=1}^{\infty}$  is a real-valued  $d$ -dimensional sequence of independently distributed random vectors with identity dispersion matrix and positive definite  $\Omega_k$ . We additionally assume the structural shocks  $\varepsilon_t$  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 mixture weights  $\alpha_{k,t} = \mathbb{P}[s_{k,t} = 1 \mid \mathcal{F}_{t-1}]$  in order to derive the probability density of the time series of interest  $y_t$ . Given a particular specification of these weights, the conditional probability density function (pdf) can be obtained by

$$f(y_t \mid \mathcal{F}_{t-1}) = \sum_{k=1}^K \alpha_{k,t} f_k(y_t \mid \mathcal{F}_{t-1}) \quad (2)$$

where  $f_k(y_t \mid \mathcal{F}_{t-1})$  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, it nests the Hamilton filter implied by the popular Markov-Switching VAR (MSVAR) model

(Krolzig 1997) with  $\alpha_{kt} = \sum_{i=1}^K p_{ik} \mathbb{1}(s_{i,t-1} = 1)$  where  $p_{ik}$  is the transition probability from state  $i$  to state  $k$ . 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  $\alpha_{k,t}$  – are encompassed as well.<sup>4</sup> 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 the weights sum up to one in each period, that is, the weights should result from the  $K$ -dimensional probability simplex  $\mathbb{S}^K = \left\{ \alpha \in \mathbb{R}_+^K \mid \sum_{i=1}^K \alpha_i = 1 \right\}$ . We achieve this by using a mapping  $h : \mathbb{R}^{K-1} \rightarrow \mathbb{S}^K$  with  $h(\tilde{\alpha}_t) = \alpha_t$  for  $\tilde{\alpha}_t \in \mathbb{R}^{K-1}$ . This ensures that the time-varying mixture weights  $\alpha_t$  sum up to one, while the dynamics of  $\tilde{\alpha}_t$  are modeled in the unrestricted domain space. We now define the parameter updating as

$$\tilde{\alpha}_t = \omega + \sum_{i=1}^p A_i S_{t-i}^{-1} \nabla_{t-i} + \sum_{i=1}^q B_i \tilde{\alpha}_{t-i} + C x_t \quad (3)$$

---

<sup>4</sup>We present and discuss the particular specification of  $\alpha_{k,t}$  in Section 3.3.

with<sup>5</sup>

$$\nabla_t = \frac{\partial \ln f(y_t | \mathcal{F}_{t-1})}{\partial \tilde{\alpha}_t} = \mathcal{J}_h(\tilde{\alpha}_t) \cdot \begin{pmatrix} \frac{f_1(y_t | \mathcal{F}_{t-1})}{f(y_t | \mathcal{F}_{t-1})} \\ \vdots \\ \frac{f_K(y_t | \mathcal{F}_{t-1})}{f(y_t | \mathcal{F}_{t-1})} \end{pmatrix} \quad (4)$$

where  $\mathcal{J}_h(\alpha_t)$  is the Jacobian of the mapping  $h$ ,  $x_t$  is an  $n$ -dimensional process of exogenous covariates measurable with respect to  $\mathcal{F}_{t-1}$ , and  $\omega \in \mathbb{R}^{k-1}$ ,  $A_i, B_i \in \mathbb{R}^{(k-1) \times (k-1)}$ , as well as  $C \in \mathbb{R}^{(k-1) \times n}$  are matrices of parameters. The updating equation (3) is a score-driven model in the sense of Creal et al. (2013) and Harvey (2013) if  $C = 0$  holds. When the latter parameter matrix is nonzero, the model is a member of the quasi-score-driven model class introduced by Blasques et al. (2023). The core feature of score-driven models is the choice of the driving innovation  $s_t = S_t^{-1} \nabla_t$  where  $S_t$  is a matrix that scales the impact of the observations on the parameter updating. A common choice is the Fisher information matrix  $\mathcal{I}_t = \mathbb{E}(\nabla_t \nabla_t^\top | \mathcal{F}_{t-1})$  or the Cholesky factor thereof to relate the scaling to the variance of the likelihood score. However, the Fisher information of observation densities cannot be derived in closed-form for our particular mixture model. Therefore, we continue with another frequent choice and set  $S_t$  equal to the identity matrix, that is,  $S_t = I_{K-1}$ . This results in the (scaled) score not having a unit variance (as in the case of a Cholesky factor) and not being adjusted for possible covariation of the scores with respect to different component weights. Despite these potential drawbacks, 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<sup>6</sup> does not crucially improve the model performance while increasing the computational burden significantly.

---

<sup>5</sup>A derivation of the score  $\nabla_t$  is provided in Appendix A.1.

<sup>6</sup>The Fisher information  $\mathcal{I}_t = \mathcal{J}_h(\alpha_t) H_t \mathcal{J}_h(\alpha_t)^\top$  can be approximated with numerical integration of

$$(H_t)_{i,j} = \int \frac{f_i(y_t | \mathcal{F}_{t-1}) f_j(y_t | \mathcal{F}_{t-1})}{f(y_t | \mathcal{F}_{t-1})} dy_t. \quad (5)$$

The score-driven 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 current observation density of the component to be updated. 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 weight updating is that it is invariant with respect to the component distribution models. Hence, the updating is not only valid for Gaussian component models (as assumed for many dynamic mixture VAR specifications) but also for other distributions.

The driving innovation term in equation (3) is accompanied by an autoregressive and a covariate term. These two features allow parsimonious modelling for persistent weight dynamics and the incorporation of external information to improve the predictive performance of the score-driven updating scheme.

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

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

with the Jacobian given by<sup>7</sup>

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

---

<sup>7</sup>A derivation of the Jacobian is provided in Appendix A.2.

We subsequently refer to the process given by the equations (1) to (7) as SMXVAR model<sup>8</sup> and to the special case where  $C = 0$  as SMVAR model. This logistic transformation is also used by Bazzi et al. (2017) to create 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 improvement in our VAR context and, consequently, decided to keep the rather simplistic logistic transformation. Additionally, given this choice of transformation function, the model also encompasses the LMVAR model of Burgard et al. (2019).

### 3.2 Two-Component Models

For simplicity, we present and discuss the properties of the SMXVAR 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 applications in Section 5. In a two-regime model, the observation density simplifies to

$$f(y_t | \mathcal{F}_{t-1}) = \frac{\exp(\tilde{\alpha}_t)}{1 + \exp(\tilde{\alpha}_t)} f_1(y_t | \mathcal{F}_{t-1}) + \frac{1}{1 + \exp(\tilde{\alpha}_t)} f_2(y_t | \mathcal{F}_{t-1}). \quad (8)$$

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

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

---

<sup>8</sup>We follow the naming convention of literature to add an X to models including external covariates (e.g., the ARMAX or the VARX model).

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

$$\tilde{\alpha}_{t+1} = \omega + a \frac{\exp(\tilde{\alpha}_t)}{(1 + \exp(\tilde{\alpha}_t))^2} \frac{f_1(y_t | \mathcal{F}_{t-1}) - f_2(y_t | \mathcal{F}_{t-1})}{f(y_t | \mathcal{F}_{t-1})} + b\tilde{\alpha}_t + cx_t. \quad (10)$$

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  $[f_1(y_t | \mathcal{F}_{t-1}) - f_2(y_t | \mathcal{F}_{t-1})] / f(y_t | \mathcal{F}_{t-1})$ . If the current observation  $y_t$  is more likely to be drawn from the first component VAR model (indicated by  $f_1(y_t | \mathcal{F}_{t-1}) > f_2(y_t | \mathcal{F}_{t-1})$  in the numerator), the latent variable  $\tilde{\alpha}_{t+1}$  is increased, given  $a > 0$ . 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 SM(X)VAR described in the previous subsections.

#### 3.3.1 Gaussian Mixture Vector Autoregression

The Gaussian Mixture Vector Autoregressive (GMVAR) 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  $p_k = \bar{p}$  for  $k = 1, \dots, K$ . For the weight updating, they stack  $\bar{p}$  lags of the time series model in a  $d\bar{p}$ -dimensional vector  $\mathbf{y}_t = \text{vec}(y_t, y_{t-1}, \dots, y_{t-\bar{p}+1})^\top$ . Hence, the regime component  $k$  features the common density

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

with unconditional mean  $\mu_k$  and covariance matrix  $\Sigma_k$ , which are functions of the parameter matrices  $\Phi_{k0}, \Phi_{k1}, \dots, \Phi_{k\bar{p}}$ , and  $\Omega_k$ . The weight updating sets the weight equal to the ratio of the observation density  $f_k(\mathbf{y}_{t-1})$  and the overall observation density  $f(\mathbf{y}_{t-1}) = \sum_{i=1}^K \alpha_K f_k(\mathbf{y}_{t-1})$  of the mixture model for each regime. That is,

$$\alpha_{k,t} = \alpha_k \frac{f_k(\mathbf{y}_{t-1})}{f(\mathbf{y}_{t-1})} \quad (12)$$

where  $\alpha = (\alpha_1, \dots, \alpha_K)^\top \in \mathbb{R}^K$  are unknown parameters. This parameterization also ensures that the dynamic weights stay within the unit simplex (i.e.,  $\sum_{k=1}^K \alpha_{k,t} = 1$ ). 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 of the dynamic weighting in the GMVAR in Eq. (12) and the SMXVAR in Eq. (4) are conceptually similar. Both approaches infer information from the ratios of the component model densities and the density of the overall mixture model. While the SMXVAR relies on conditional observation densities, the common unconditional density of the  $\bar{p}$  most recent observations pins down the dynamic weights in the GMVAR. While the GMVAR is theoretically more appealing as the resulting time series is a Markovian process, the SMXVAR offers more flexibility in describing the data. First, the autoregressive term in the score-driven updating in Eq. (4) allows parsimonious modeling of persistence in the mixture weights. A GMVAR could accommodate such persistent mixture dynamics only by increasing the number of lags  $\bar{p}$ , which crucially increases the number of parameters to be estimated. Second, the SMXVAR can incorporate external predictor information and does not have to rely on recent observations of the time series variables only. Third, our model explicitly allows for non-Gaussian innovations in the component VARs.

### 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

$$\alpha_{k,t} = \frac{\exp\{\gamma'_k x_{t-1}\}}{\sum_{j=1}^K \exp\{\gamma'_j x_{t-1}\}} \quad (13)$$

with  $x_{t-1}$  being a vector of  $\mathcal{F}_{t-1}$ -measurable predictors and  $\gamma_k$  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. (2023) use the component ratio from Eq. (12) of the GMVAR as predictor of the weight dynamics. Hence, the component ratio driving the SMXVAR mixture update is also used as driving variable. Both, the LMVAR and the SMXVAR, use a logistic transformation to map weights onto the unit interval. However, the SMXVAR maps an unrestricted internal process that may use external predictors, whereas the LMVAR directly ties the weights to the predictors using an internal multinomial logit regression. More concretely, the LMVAR is a special case of the SMXVAR if the coefficients of scores and prior weights in the updating scheme are zero (i.e.,  $A_i = B_i = 0$ ). Hence, the SMXVAR can be a worthwhile alternative to the LMVAR, especially when the predictors in  $x_t$  are imperfect (for simulation evidence, see Section 4.2.2).

## 3.4 Estimation and Inference

We estimate the vector  $\theta_k = (\Phi'_0, \text{vec}(\Phi_{k1})', \dots, \text{vec}(\Phi_{k,p_k})', \text{vech}(\Omega_k)')$ <sup>9</sup> containing the VAR parameters for each component submodel. Furthermore, we estimate the parameters characterizing the mixture weight updating  $\theta_s = (\omega', \text{vec}(A_1)', \dots, \text{vec}(A_p)', \text{vec}(B_1)', \dots, \text{vec}(B_p)')$  as well as possible additional parameters characterizing the distribution of  $\varphi$  collected in

---

<sup>9</sup> $\text{vec}(A)$  denotes the vectorization that stacks all columns of the matrix  $A$  into a vector and  $\text{vech}(A)$  the half-vectorization that only stacks the upper triangular part of  $A$ .

vector  $\theta$ . In summary, this yields the complete parameter vector  $\vartheta = (\theta, \theta_1, \dots, \theta_k, \theta_s)$  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 and is given by

$$\ln \mathcal{L} = \ln \prod_{t=1}^T f(y_t | \mathcal{F}_{t-1}) = \sum_{t=1}^T \ln \sum_{k=1}^K \alpha_{k,t} f_k(y_t | \mathcal{F}_{t-1}) \quad (14)$$

$$= \sum_{t=1}^T \ln \sum_{k=1}^K \alpha_{k,t} \varphi \left( \Omega_k^{-\frac{1}{2}} (y_t - \Phi_{k0} - \Phi_{k1} y_{t-1} - \dots - \Phi_{kp_k} y_{t-p_k}) \right) \quad (15)$$

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  $\vartheta$  stacks all the static parameters of the model, standard asymptotic theory for ML estimators suggests that, under some regularity conditions, the following holds:

$$\sqrt{T} \left( \hat{\vartheta} - \vartheta \right) \xrightarrow{d} \mathcal{N} \left( 0, \mathcal{I}^{-1}(\vartheta) \right) \quad (16)$$

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

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.<sup>10</sup> Another issue, as noted by Krolzig (1997), is an identification problem that is superficially caused by the interchangeability of state labels. We identify states in the following two-state application by the unconditional mixture weights according to  $\bar{\alpha}_1 > \bar{\alpha}_2$ . For deriving standard errors, we numerically compute the Fisher information matrix with a finite difference scheme.

---

<sup>10</sup>For this reason, we employ the OptQuest Nonlinear Programming (OQNLP) multistart heuristic algorithm of Ugray et al. (2007). The heuristic comprises a global phase and a local phase. The global phase involves a preliminary scatter search, in which trial points are generated and subsequently filtered to yield a smaller subset of candidate starting points. In the local phase, local gradient-based NLP solvers (the interior-point algorithm of Byrd et al. (1999) in our case) are employed to identify a local optimum from these starting points.

### 3.5 Impulse Responses

A popular VAR-based tool to investigate the impact of shocks, especially in empirical macroeconomics, is the impulse response analysis. Given a fitted and identified VAR model as well as initial values, responses to 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

$$GI(n, \xi_t, \mathcal{F}_{t-1}) = \mathbb{E}(y_{t+n} | \xi_t, \mathcal{F}_{t-1}) - \mathbb{E}(y_{t+n} | \mathcal{F}_{t-1}) \quad (17)$$

where  $n$  is the number of periods ahead and  $\xi_t$  is a structural shock that occurs in period  $t$  as realization of the error term  $\varepsilon_t$ .<sup>11</sup> Hence, Eq. (17) defines the impulse response as difference between the  $n$ -period ahead prediction given the shock and the corresponding prediction in absence of the shock.

We follow a Monte Carlo approach to derive the generalized impulse responses to the shock  $\xi_t$  in (17). 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  $(\varepsilon_{t+n})_{n=0}^N$  are randomly drawn. Furthermore, a second series  $(y_{t+n}(\xi_t, \mathcal{F}_{t-1}))_{n=0}^N$  is constructed for which the same random shocks are used but  $\varepsilon_t = \xi_t$  is imposed. Then, we can use the means of these artificially drawn series to approximate the expectations in (17) and compute GI by taking differences. The algorithm used in the empirical applications is described in more detail in Appendix C.3.

## 4 Monte Carlo Study

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

---

<sup>11</sup>Virolainen (2020) uses the same concept to derive impulse responses for a structural version of the GMVAR model.

## 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:

$$\begin{aligned}
 y_t = & s_{1,t} \left( \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0.5 & 0.2 \\ 0.2 & 0.5 \end{pmatrix} y_{t-1} + \begin{pmatrix} 0.5 & 0 \\ 0.2 & 0.5 \end{pmatrix} \epsilon_t \right) \\
 & + s_{2,t} \left( \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.6 & 0 \\ 0.4 & 0.6 \end{pmatrix} y_{t-1} + \begin{pmatrix} 0.2 & 0 \\ 0.8 & 0.2 \end{pmatrix} \epsilon_t \right).
 \end{aligned} \tag{18}$$

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  $T \in \{200, 500, 1000\}$  and then estimate the parameter vector  $\vartheta$  for a Gaussian<sup>12</sup> SMVAR with lags  $p = q = 1$  and  $K = 2$  using likelihood optimization. This means that the weight updating is performed according to Eq. (10). We also fit an MVAR, an MSVAR, and a GMVAR as benchmarks to evaluate the performance of our SMVAR model. In addition, we also estimate the DGP V using SMXVAR and LMVAR models, which are able to use external predictor information in  $x_t$ .

---

<sup>12</sup>We restrict the simulation study to the Gaussian case in order to preserve comparability with the benchmark models that require Gaussianity, even though the SMVAR is not restricted to this case.

**Table 1:** Mixture Processes of the Simulation Study

| DGP |                             | Mixture Weights ( $\alpha_t$ )  |
|-----|-----------------------------|---|
| I   | Constant                    | 0.5   |
| II  | Break                       | $0.8 \cdot \mathbf{1}(t < \frac{T}{2}) + 0.2 \cdot \mathbf{1}(t \geq \frac{T}{2})$                                  |
| III | Cycle                       | $0.5 + 0.45 \cdot \sin(4\pi t/T)$   |
| IV  | Markov-Switching (MS)       | $0.95 \cdot \mathbf{1}(s_{t-1} = 1) + 0.10 \cdot \mathbf{1}(s_{t-1} = 2)$   |
| V   | Logit Auto-Regressive (LAR) | $\alpha_t = 1/(1 + \exp\{-0.5x_t\})$<br>$x_t = 0.97x_{t-1} + u_t, \quad u_t \stackrel{iid}{\sim} \mathcal{N}(0, 1)$ |

## 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 the SMXVAR and LMVAR in the presence of external factor-driven mixture dynamics.

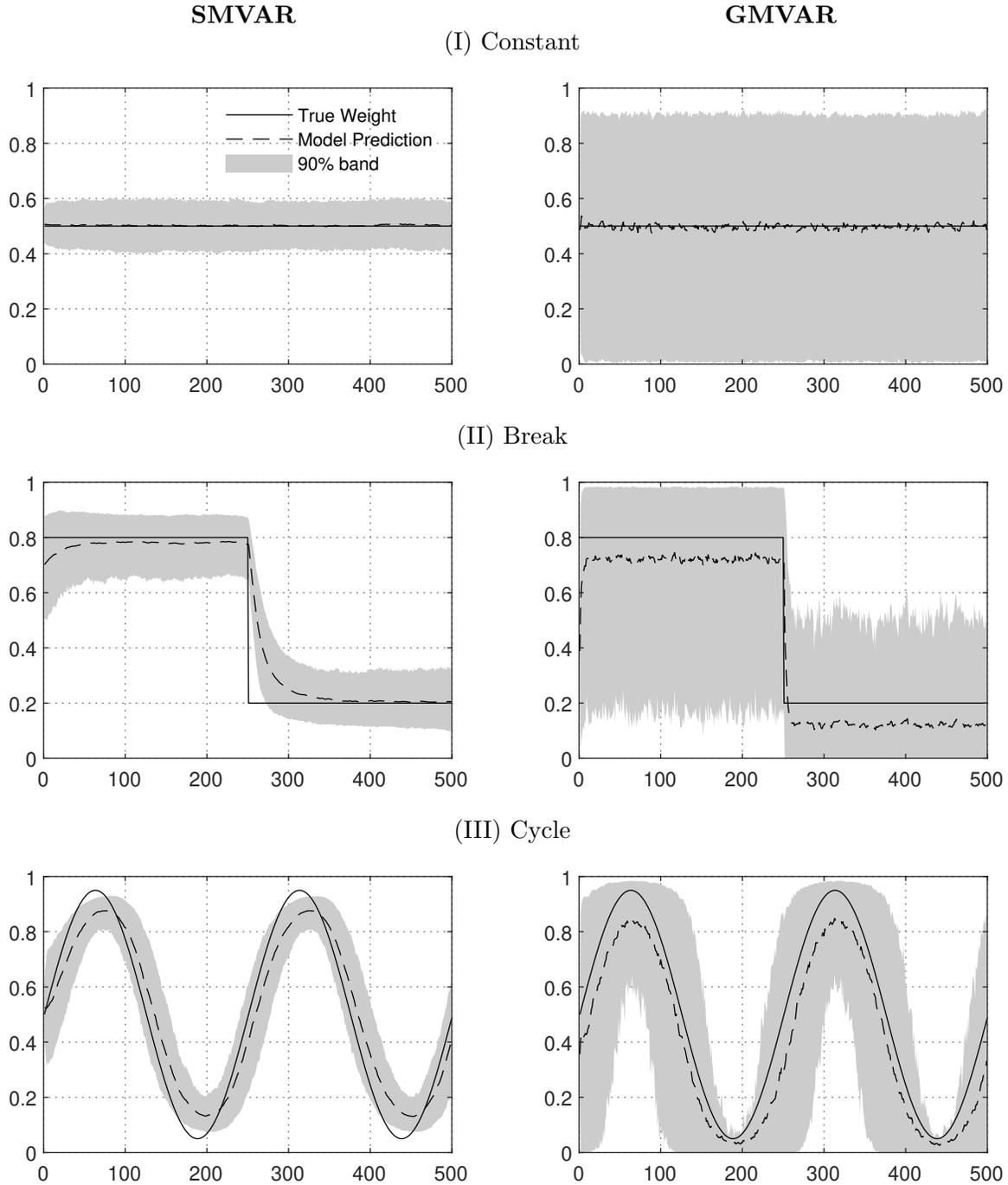
### 4.2.1 Comparison of SMVAR and GMVAR

Figure 1 shows the average SMVAR and GMVAR estimates of  $\alpha_t$  fitted to 1000 Monte Carlo replications and  $T = 500$  alongside the true weights for DGPs I, II, and III.<sup>13</sup> The shaded areas indicate 90% 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 requires a certain time period 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.

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

<sup>13</sup>Note that a similar graphical representation of average estimates is not meaningful for the stochastic DGPs (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**



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% bands.

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  $f_1(y_{t-1})/f(y_{t-1})$  is constant over time. Hence, the DGP I with  $\alpha_{1,t} \equiv \alpha_k = 0.5$  is not nested by the GMVAR. Here, the SMVAR benefits from the additional constant parameter  $\omega$  that allows an incorporation of any constant level in the static case, that is, if  $a$  and  $b$  are zero. Indeed, we find that the learning and persistence parameters are estimated close to zero with  $\hat{a} = 0.0328$  and  $\hat{b} = 0.0996$  on average. The estimate for  $\hat{\omega}$  is 0.078 on average, which translates into an average unconditional weight of  $\hat{\lambda} = 0.5019$ . The same nesting problem is visible in the break process (II) that is estimated with huge bands. The two constant levels cannot be captured correctly – even on average – because the single parameter  $\alpha_1$  has to serve as average weight and changing probability of 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 faster than by the SMVAR. The cycles are captured quite well by the GMVAR, although the 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 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 estimates of the mixture weight  $\alpha_{1,t}$  from MVAR, MSVAR, GMVAR, and SMVAR models with 1000 Monte Carlo replications and different sample sizes  $T$ . Table B.1 in Appendix B states 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

**Table 2:** Average Estimation Error Comparison

| DGP        | $T$  | MSE          |              |       |              | MAE          |              |       |              |
|------------|------|--------------|--------------|-------|--------------|--------------|--------------|-------|--------------|
|            |      | MVAR         | MSVAR        | GMVAR | SMVAR        | MVAR         | MSVAR        | GMVAR | SMVAR        |
| I Constant | 200  | <b>0.006</b> | 0.010        | 0.128 | 0.010        | <b>0.062</b> | 0.076        | 0.321 | 0.076        |
|            | 500  | <b>0.002</b> | 0.003        | 0.108 | 0.004        | <b>0.037</b> | 0.045        | 0.291 | 0.051        |
|            | 1000 | <b>0.001</b> | 0.002        | 0.096 | 0.003        | <b>0.025</b> | 0.031        | 0.271 | 0.043        |
| II Break   | 200  | 0.108        | 0.080        | 0.088 | <b>0.027</b> | 0.301        | 0.230        | 0.231 | <b>0.117</b> |
|            | 500  | 0.103        | 0.072        | 0.080 | <b>0.018</b> | 0.300        | 0.221        | 0.218 | <b>0.095</b> |
|            | 1000 | 0.102        | 0.070        | 0.078 | <b>0.012</b> | 0.300        | 0.218        | 0.215 | <b>0.076</b> |
| III Cycle  | 200  | 0.122        | 0.079        | 0.090 | <b>0.034</b> | 0.302        | 0.222        | 0.225 | <b>0.141</b> |
|            | 500  | 0.116        | 0.072        | 0.083 | <b>0.018</b> | 0.297        | 0.213        | 0.214 | <b>0.104</b> |
|            | 1000 | 0.116        | 0.069        | 0.080 | <b>0.013</b> | 0.297        | 0.209        | 0.210 | <b>0.088</b> |
| IV MS      | 200  | 0.194        | <b>0.067</b> | 0.104 | 0.070        | 0.409        | <b>0.121</b> | 0.176 | 0.157        |
|            | 500  | 0.199        | <b>0.066</b> | 0.104 | 0.067        | 0.431        | <b>0.116</b> | 0.175 | 0.139        |
|            | 1000 | 0.204        | 0.067        | 0.104 | <b>0.066</b> | 0.442        | <b>0.113</b> | 0.174 | 0.132        |
| V LAR      | 200  | 0.087        | 0.065        | 0.103 | <b>0.041</b> | 0.244        | 0.202        | 0.243 | <b>0.159</b> |
|            | 500  | 0.098        | 0.066        | 0.095 | <b>0.039</b> | 0.267        | 0.208        | 0.232 | <b>0.155</b> |
|            | 1000 | 0.104        | 0.068        | 0.093 | <b>0.038</b> | 0.277        | 0.211        | 0.229 | <b>0.155</b> |

Average mean squared error (MSE) and average mean absolute error (MAE) for the estimate of the mixture weight  $\alpha_{1,t}$  from MVAR, MSVAR, GMVAR, and SMVAR models with 1000 Monte Carlo replications and different sample sizes  $T$ . Bold numbers depict the best-performing models.

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 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 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:** Average Prediction Error Comparison

| DGP        | $T$  | MSPE         |              |       |              | MAPE         |              |       |              |
|------------|------|--------------|--------------|-------|--------------|--------------|--------------|-------|--------------|
|            |      | MVAR         | MSVAR        | GMVAR | SMVAR        | MVAR         | MSVAR        | GMVAR | SMVAR        |
| I Constant | 200  | <b>0.006</b> | 0.010        | 0.128 | <b>0.020</b> | 0.062        | 0.076        | 0.321 | 0.109        |
|            | 500  | <b>0.002</b> | 0.003        | 0.110 | <b>0.007</b> | 0.037        | 0.045        | 0.293 | 0.064        |
|            | 1000 | <b>0.001</b> | 0.002        | 0.098 | <b>0.003</b> | 0.025        | 0.031        | 0.276 | 0.046        |
| II Break   | 200  | 0.063        | 0.068        | 0.079 | <b>0.021</b> | 0.227        | 0.209        | 0.235 | <b>0.112</b> |
|            | 500  | 0.045        | 0.060        | 0.062 | <b>0.012</b> | 0.203        | 0.197        | 0.212 | <b>0.091</b> |
|            | 1000 | 0.042        | 0.049        | 0.052 | <b>0.009</b> | 0.200        | 0.172        | 0.203 | <b>0.073</b> |
| III Cycle  | 200  | <b>0.021</b> | 0.063        | 0.164 | 0.045        | <b>0.121</b> | 0.218        | 0.378 | 0.180        |
|            | 500  | <b>0.015</b> | 0.058        | 0.146 | 0.029        | <b>0.108</b> | 0.213        | 0.349 | 0.140        |
|            | 1000 | <b>0.014</b> | 0.052        | 0.140 | 0.018        | 0.113        | 0.204        | 0.341 | <b>0.109</b> |
| IV MS      | 200  | 0.199        | <b>0.035</b> | 0.068 | 0.044        | 0.415        | <b>0.081</b> | 0.137 | 0.122        |
|            | 500  | 0.200        | <b>0.029</b> | 0.053 | 0.033        | 0.432        | <b>0.072</b> | 0.125 | 0.100        |
|            | 1000 | 0.206        | <b>0.031</b> | 0.065 | 0.033        | 0.444        | <b>0.073</b> | 0.135 | 0.093        |
| V LAR      | 200  | 0.108        | 0.079        | 0.090 | <b>0.041</b> | 0.273        | 0.221        | 0.225 | <b>0.159</b> |
|            | 500  | 0.106        | 0.072        | 0.088 | <b>0.035</b> | 0.279        | 0.220        | 0.221 | <b>0.145</b> |
|            | 1000 | 0.111        | 0.075        | 0.093 | <b>0.035</b> | 0.289        | 0.223        | 0.229 | <b>0.148</b> |

Mean squared prediction error (MSPE) and mean absolute prediction error (MAPE) for the out-of-sample prediction of the mixture weight  $\alpha_{1,T+1}$  from MVAR, MSVAR, GMVAR, and SMVAR models with 1000 Monte Carlo replications and different sample sizes  $T$ . Bold numbers depict the best-performing models.

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 MVAR, MSVAR, GMVAR, and SMVAR models with 1000 Monte Carlo replications and different sample sizes  $T$ . The results show that the performance of the SMVAR in comparison to the GMVAR also holds out-of-sample.<sup>14</sup>

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

<sup>14</sup>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.

accommodate a wide range of practically relevant DGPs which the SMVAR is able to handle adequately by adding a constant and an autoregressive component to the mixture weight updating scheme.

#### 4.2.2 Comparison of SMXVAR and LMVAR

The LMVAR of Burgard et al. (2019) uses covariates to explain the mixture weight dynamics. This is particularly 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. The SMXVAR is able to consider external predictors for the weights, while the general updating is still driven by a GAS process as in the SMVAR. In the following, we examine the case where the explanatory value of the external factor is a priori uncertain. We investigate whether the score-driven dynamics of the SMXVAR can compensate for imperfectly measured predictors provided to the model.

Again, we simulate from the VAR model given by Eq. (18) now using DGP V as mixture process. We also use 1000 Monte Carlo replications of different sample sizes  $T \in \{200, 500, 1000\}$ . Note that the dynamics of  $\alpha_t$  follow an LMVAR model as described by Eq. (13) with  $K = 2$ . Hence, if we estimate an LMVAR given the true covariate  $x_t$ , it will trivially outperform the other models. The interesting question is what happens if we provide the models with a noisy signal  $\tilde{x}_t$  of  $x_t$ . We create this signal by

$$\tilde{x}_t = \rho x_t + (1 - \rho)\tilde{u}_t, \quad \tilde{u}_t \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2) \quad (19)$$

where  $\sigma^2 = 16.9205$  is equal to the unconditional variance of the process  $x_t$  as defined in Table 1. 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 estimates of the mixture weight  $\alpha_{1,t}$  from an SMXVAR and an LMVAR model. We clearly

**Table 4:** Comparison of Average Estimation Errors: SMXVAR vs. LMVAR

| $T$      | SMXVAR       |              |              |              |              |       | LMVAR      |       |       |       |       |              |
|----------|--------------|--------------|--------------|--------------|--------------|-------|------------|-------|-------|-------|-------|--------------|
|          | $\rho = 0$   | 0.2          | 0.4          | 0.6          | 0.8          | 1     | $\rho = 0$ | 0.2   | 0.4   | 0.6   | 0.8   | 1            |
| (a) MSE  |              |              |              |              |              |       |            |       |       |       |       |              |
| 200      | <b>0.060</b> | <b>0.044</b> | <b>0.033</b> | <b>0.022</b> | <b>0.012</b> | 0.009 | 0.095      | 0.088 | 0.069 | 0.036 | 0.012 | <b>0.005</b> |
| 500      | <b>0.046</b> | <b>0.036</b> | <b>0.027</b> | <b>0.016</b> | <b>0.007</b> | 0.003 | 0.100      | 0.094 | 0.071 | 0.035 | 0.008 | <b>0.002</b> |
| 1000     | <b>0.040</b> | <b>0.035</b> | <b>0.027</b> | <b>0.015</b> | <b>0.006</b> | 0.001 | 0.105      | 0.099 | 0.074 | 0.035 | 0.008 | <b>0.001</b> |
| (b) MAE  |              |              |              |              |              |       |            |       |       |       |       |              |
| 200      | <b>0.191</b> | <b>0.161</b> | <b>0.137</b> | <b>0.109</b> | <b>0.080</b> | 0.065 | 0.256      | 0.246 | 0.212 | 0.146 | 0.080 | <b>0.052</b> |
| 500      | <b>0.166</b> | <b>0.148</b> | <b>0.126</b> | <b>0.094</b> | <b>0.061</b> | 0.036 | 0.270      | 0.261 | 0.221 | 0.144 | 0.067 | <b>0.029</b> |
| 1000     | <b>0.158</b> | <b>0.147</b> | <b>0.124</b> | <b>0.091</b> | <b>0.055</b> | 0.027 | 0.279      | 0.270 | 0.227 | 0.145 | 0.064 | <b>0.021</b> |
| (c) MSPE |              |              |              |              |              |       |            |       |       |       |       |              |
| 200      | <b>0.070</b> | <b>0.049</b> | <b>0.034</b> | <b>0.023</b> | <b>0.013</b> | 0.009 | 0.113      | 0.107 | 0.083 | 0.042 | 0.013 | <b>0.005</b> |
| 500      | <b>0.047</b> | <b>0.034</b> | <b>0.026</b> | <b>0.015</b> | <b>0.007</b> | 0.003 | 0.105      | 0.099 | 0.075 | 0.038 | 0.010 | <b>0.002</b> |
| 1000     | <b>0.043</b> | <b>0.036</b> | <b>0.026</b> | <b>0.014</b> | <b>0.006</b> | 0.002 | 0.107      | 0.101 | 0.076 | 0.036 | 0.008 | <b>0.001</b> |
| (d) MAPE |              |              |              |              |              |       |            |       |       |       |       |              |
| 200      | <b>0.205</b> | <b>0.170</b> | <b>0.140</b> | <b>0.109</b> | <b>0.081</b> | 0.066 | 0.282      | 0.274 | 0.236 | 0.159 | 0.084 | <b>0.054</b> |
| 500      | <b>0.165</b> | <b>0.144</b> | <b>0.122</b> | <b>0.093</b> | <b>0.063</b> | 0.039 | 0.275      | 0.264 | 0.226 | 0.150 | 0.072 | <b>0.030</b> |
| 1000     | <b>0.162</b> | <b>0.148</b> | <b>0.122</b> | <b>0.088</b> | <b>0.055</b> | 0.029 | 0.282      | 0.272 | 0.230 | 0.149 | 0.066 | <b>0.021</b> |

Average mean squared error (MSE), average mean absolute error (MAE), mean squared prediction error (MSPE), and average mean absolute prediction error (MAPE) for the estimate of the mixture weight  $\alpha_{1,t}$  from an SMXVAR and an LMVAR model with 1000 Monte Carlo replications and different sample sizes  $T$ . The parameter  $\rho$  indicates the informational content of the used covariate signal. Bold numbers indicate which of the two models performed better.

see that the LMVAR using the true covariate outperforms all other models (as expected). However, when the signal is imperfect, the SMXVAR almost always outperforms the LMVAR. The only exceptions are the two in-sample criteria MSE and MAE when  $\rho = 0.8$  and  $T = 200$ , where the values are almost the same for both models. Strikingly, the performance of the LMVAR decreases much faster for decreasing  $\rho$  than it does for the SMXVAR. In particular, for  $\rho \leq 0.2$ , the LMVAR is not able to outperform the static MVAR benchmark (see DGP V, columns MVAR in Table 2) while the SMVAR still shows an adequate performance for  $\rho = 0$  because of its score-driven term.

Also noteworthy is that the SMVAR without external information from  $x_t$  (see Table 2) outperforms the LMVAR if the signal is noisy with  $\rho \in \{0, 0.2, 0.4\}$ . The break-even point at which both methods, SMVAR and LMVAR, are equal may be slightly above a  $\rho$  of 0.5. Hence, more than 50% of the variation in the signal must come from the true covariate to justify the use of LMVAR in this setting. As demonstrated throughout the literature, noise ratios of 50% or more are not uncommon for many economic and, in particular, financial variables. A similar result is found for the out-of-sample measures in panels (c) and (d) of Table 4, where the break-even value of noise improves slightly in favor of the SMVAR (see Table 3).

The main insight from this simulation exercise is that the SMXVAR is a worthwhile alternative to the LMVAR and should be prioritized when external predictors are of questionable quality. The key advantage is that the score-driven part of the SMXVAR update can compensate for imperfections of the covariates.

## 5 Empirical Applications

We demonstrate the applicability of SM(X)VAR models through two empirical applications. Application 1 focuses on model performance relative to other dynamic MVAR specifications and weight predictability in stock and bond return dynamics. Application 2 particularly discusses the impact of shocks on regime probabilities of macro-financial linkages.

### 5.1 Application 1: Stock and Bond Return Dynamics

Many authors have documented regime shifts in the relation of stock and bond returns and model these with Markov-Switching processes (see, for example, Guidolin and Ono 2006; Guidolin and Timmermann 2006; Kole and van Dijk 2023). We further investigate the regime dynamics of this relationship with a score-driven mixture VAR.

### 5.1.1 SMVAR Results

We fit a two-state Gaussian SMVAR model using data from the Center for Research on Security Prices (CRSP). We use value-weighted equity returns (including dividends) from the NYSE, NASDAQ, and AMXE stock exchanges as stock market index. The bond index return is based on 10-year to maturity US government bonds. We derive excess returns of the two markets by subtracting the 1-month T-Bill rate. As supported by both information criteria, we use  $K = 2$  components. Considering lag lengths of up to 4 (i.e., 25 combinations of  $p_1$  and  $p_2$ ), both AIC and BIC favor one lag in both VAR components.

**Table 5:** Estimation Results for Stock and Bond Return Dynamics

| <b>(a) VAR Parameters</b>           |                         |                          |                         |                          |
|-------------------------------------|-------------------------|--------------------------|-------------------------|--------------------------|
|                                     | Normal Regime           |                          | Crisis Regime           |                          |
|                                     | <i>Bond<sub>t</sub></i> | <i>Stock<sub>t</sub></i> | <i>Bond<sub>t</sub></i> | <i>Stock<sub>t</sub></i> |
| <i>Bond<sub>t-1</sub></i>           | 0.063<br>(0.048)        | -0.053<br>(0.022)        | 0.091<br>(0.062)        | -0.086<br>(0.031)        |
| <i>Stock<sub>t-1</sub></i>          | 0.162<br>(0.112)        | -0.048<br>(0.053)        | 0.272<br>(0.130)        | 0.079<br>(0.065)         |
| <i>const</i>                        | 0.069<br>(0.075)        | 1.247<br>(0.183)         | 0.273<br>(0.166)        | -0.354<br>(0.358)        |
| <b>(b) (Co-)Variance Parameters</b> |                         |                          |                         |                          |
|                                     | Normal Regime           |                          | Crisis Regime           |                          |
| $\Omega_{1,1}$                      | 1.724<br>(0.200)        |                          | 8.108<br>(0.806)        |                          |
| $\Omega_{1,2}$                      | 0.040<br>(0.261)        |                          | 2.125<br>(0.950)        |                          |
| $\Omega_{2,2}$                      | 8.149<br>(0.885)        |                          | 32.977<br>(3.099)       |                          |
| <b>(c) GAS Parameters</b>           |                         |                          |                         |                          |
|                                     | $\bar{\alpha}$          | $a$                      | $b$                     |                          |
|                                     | 0.423<br>(0.444)        | 1.706<br>(0.404)         | 0.952<br>(0.018)        |                          |

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 given in parentheses.

Table 5 shows the estimation results for the SMVAR model (with standard errors in parentheses). The first component features moderate variances ( $\Omega_{1,1}$  and  $\Omega_{2,2}$ ) and an insignificant covariance of the two return series ( $\Omega_{1,2}$ ). The only significant coefficient in the model represents a negative impact of lagged bond returns on stock returns. The second component shows much higher variances and a significantly positive covariance. Hence, this component seems to be particularly influential in times of crisis when volatility increases and stock and bond returns typically move together. In this high volatility component, the negative spillovers from bonds to stocks are larger and bond returns are also positively affected by lagged stock returns.

Turning to the weight dynamics in panel (c) of Table 5, it has to be noted that we estimate a slightly re-parameterized internal updating process for the unrestricted mixture weights, which is given by

$$\tilde{\alpha}_{t+1} = \bar{\alpha} + a \frac{\exp(\tilde{\alpha}_t)}{(1 + \exp(\tilde{\alpha}_t))^2} \frac{p_1(y_t | \mathcal{F}_{t-1}) - p_2(y_t | \mathcal{F}_{t-1})}{p(y_t | \mathcal{F}_{t-1})} + b (\tilde{\alpha}_t - \bar{\alpha}), \quad (20)$$

where  $\bar{\alpha} = \omega/(1 - b)$  is the unconditional unrestricted mixture weight in case the process is stationary. Accordingly, the estimate for the unconditional restricted mixture weight  $\bar{\alpha}$  is given by 0.604,<sup>15</sup> implying that the first component has a slightly larger weight on average. The coefficient for the weight update  $a$  is positive and significant, indicating that the scaled observation density evaluated at the current observation is informative for detecting changes in future mixture weights. The estimate of 0.952 for the autoregressive parameter  $b$  suggests that the weights are persistent.

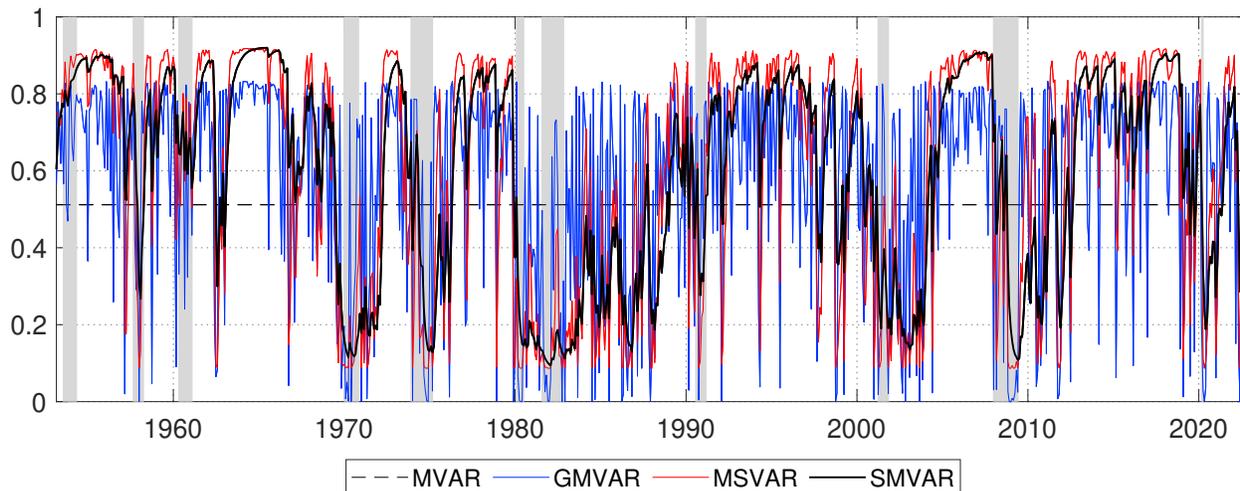
Figure 5 shows the evolution of the transformed mixture weights  $\alpha_t = \exp(\tilde{\alpha}_t)/[1 + \exp(\tilde{\alpha}_t)]$  over time based on our proposed SMVAR approach as well as MVAR, GMVAR, and MSVAR models as benchmarks. We see that the high volatility component gains importance ( $\alpha_t$  falls) in particular during NBER recessions (indicated by the shaded areas). Most striking is the sharp drop of  $\alpha_t$  at the beginning of the global financial crisis in 2008 and the gradual

---

<sup>15</sup> $\bar{\alpha} = \exp(\bar{\omega})/(1 + \bar{\omega}) = \exp(0.423)/(1 + \exp(0.423)) = 0.604$

recovery thereafter. This is in line with Guidolin and Timmermann (2006) who document the existence of a recovery state. Especially for the SMVAR, the weight dynamics appear to be slow-moving in cycles, best seen in the 1980s and 1990s. This speaks in favor of a dynamic mixture interpretation instead of a setting with clearly separated regimes. When comparing the three different dynamic approaches, it is noticeable that the GMVAR is extremely volatile. This is particularly true in the 1980s, when the weight slowly transitions between the components in the SMVAR and MSVAR. Again, we see that the GMVAR, due to its Markovian structure, is better suited to model the switching between states and is hardly able to capture slow-moving weight dynamics.

**Figure 5:** Development of  $\alpha_t$  over Time for Stock and Bond Return Dynamics



Probability of being in the normal state as indicated by the estimated mixture weights  $\alpha_t$  based on four different approaches. Shaded areas indicate NBER recessions.

### 5.1.2 Further Comparison to Alternative MVAR Models

Next, we compute the goodness-of-fit statistics and mean prediction errors after fitting different MVAR models to the same data set. These are shown in Table 6. We find that both information criteria (AIC and BIC) favor the SMVAR models. Hence, the richer parameterization of the SMVAR seems to be rewarded in terms of model fit. The MAPE and MSPE measures are based on implied return predictions instead of the unobservable mix-

**Table 6:** Likelihood-based Goodness-of-Fit and Mean Prediction Errors

| Model  | # $\vartheta$ | $L$     | AIC           | BIC           | MAPE          |               | MSPE          |               |
|--------|---------------|---------|---------------|---------------|---------------|---------------|---------------|---------------|
|        |               |         |               |               | Stocks        | Bonds         | Stocks        | Bonds         |
| MVAR   | 19            | -4157.5 | 8353.1        | 8443.0        | 1.5854        | <b>3.3910</b> | 4.5429        | <b>19.940</b> |
| MSVAR  | 20            | -4116.6 | 8273.3        | 8367.9        | 1.5846        | 3.4135        | 4.5317        | 20.159        |
| GMVAR  | 19            | -4142.8 | 8323.5        | 8413.4        | <b>1.5836</b> | 3.4050        | <b>4.5228</b> | 19.974        |
| SMVAR  | 21            | -4111.1 | 8264.2        | <b>8363.6</b> | 1.5848        | 3.4109        | 4.5348        | 20.207        |
| tSMVAR | 22            | -4110.0 | <b>8264.0</b> | 8368.1        | 1.5840        | 3.4081        | 4.5357        | 20.236        |

Column # $\vartheta$  shows the number of estimated model parameters,  $L$  the log-likelihood, AIC the Akaike Information Criterion, BIC the Bayesian Information Criterion, MSPE the mean squared prediction error, and MAPE the mean absolute prediction error. Boldface entries indicate the best model according to the respective information or prediction error criterion.

ture weights. Therefore, it is not surprising that both measures are similar for all models due to the large variation in returns that does not come from the mixture weight dynamics. Nevertheless, the prediction errors of the dynamic approaches are clearly smaller than that of the static MVAR benchmark. This is another indication of the SMVAR's appeal in cases where the actual weight process is uncertain.

As noted above, the SMVAR allows for different component distributions. Hence, we can easily generalize the model to incorporate Student's t-distributions, which we denote as tSMVAR.<sup>16</sup> This only leads to a minor improvement over the Gaussian case in terms of log-likelihood and AIC and to a decrease in BIC. This result most likely comes from the large estimated degree of freedom parameter ( $\nu = 22.854$ ) and the fact that finite mixtures of normal distributions are themselves capable of producing leptokurtic shapes, even though asymptotically the tails behave in a Gaussian manner (McLachlan and Peel 2004; Haas and Pigorsch 2009).

<sup>16</sup>Estimation results of the tSMVAR model are reported in Appendix C.2.

### 5.1.3 External Predictors Using SMXVAR

The SMXVAR offers the opportunity to specify external predictors for the dynamic weight updating. We consider several macroeconomic and financial variables that have a reported connection to bond and stock return dynamics: (i) the 3-Month T-Bill yield, (ii) the logarithm of the CRSP dividend yield, (iii) the term spread computed as difference between the yields of the 10-Year T-Bond and the 3-Month T-Bill, (iv) the default spread defined as differential yield on Moody’s Bbb and Aaa seasoned corporate bonds of the same maturity, and (v) the seasonally adjusted growth rate of industrial production from the FRED database.

Table 7 shows the GAS parameter estimates and goodness-of-fit statistics of SMXVAR model specifications using the aforementioned predictors. We see that only the 3-Month T-Bill yield (column 2) and the term spread (column 4) significantly predict the weight changes in the relationship between bond and stock returns. Whereas the short-term interest rate increases the weight of the more volatile regime, the term spread contributes to a decrease of this weight. The qualitative impact of both variables remains stable in a model using all predictors simultaneously (column 7). Remarkably, the learning rate parameter  $a$  of the GAS dynamics also remains stable when adding additional external predictors. Hence, the autoregressive score dynamics still contributes most to filtering the dynamic mixture weights.

## 5.2 Application 2: Macro-Financial Linkages

Adrian et al. (2021) document a pronounced multimodality in the conditional distribution of economic growth, especially in times of tight financial conditions. Mixture models are a typical choice for modeling such multimodalities, and the time-varying weight in the SMVAR allows to accommodate the dynamic linkages to financial conditions.

**Table 7:** SMXVAR GAS-Parameters for Stock and Bond Return Dynamics

|                    | (1)              | (2)               | (3)               | (4)               | (5)               | (6)              | (7)               |
|--------------------|------------------|-------------------|-------------------|-------------------|-------------------|------------------|-------------------|
| $\bar{\alpha}$     | 0.423<br>(0.444) | 2.005<br>(0.643)  | 0.640<br>(1.063)  | 0.663<br>(0.757)  | 2.451<br>(0.917)  | 0.367<br>(0.525) | 1.349<br>(1.089)  |
| $a$                | 1.706<br>(0.404) | 1.543<br>(0.389)  | 1.696<br>(0.404)  | 1.567<br>(0.383)  | 1.717<br>(0.440)  | 1.701<br>(0.406) | 1.323<br>(0.471)  |
| $b$                | 0.952<br>(0.018) | 0.940<br>(0.021)  | 0.953<br>(0.018)  | 0.961<br>(0.015)  | 0.907<br>(0.048)  | 0.951<br>(0.021) | 0.901<br>(0.058)  |
| 3M T-Bill Yield    | –                | –0.025<br>(0.010) | –                 | –                 | –                 | –                | –0.037<br>(0.024) |
| Log Dividend Yield | –                | –                 | –0.011<br>(0.048) | –                 | –                 | –                | 0.209<br>(0.138)  |
| Term Spread        | –                | –                 | –                 | –0.036<br>(0.018) | –                 | –                | 0.048<br>(0.037)  |
| Default Spread     | –                | –                 | –                 | –                 | –0.201<br>(0.158) | –                | –0.244<br>(0.177) |
| Ind. Prod. Growth  | –                | –                 | –                 | –                 | –                 | 0.013<br>(0.072) | 0.047<br>(0.102)  |
| $L$                | –4111.1          | –4105.9           | –4111.1           | –4109.3           | –4109.2           | –4111.1          | –4100.6           |
| AIC                | 8264.2           | 8255.7            | 8266.1            | 8262.6            | 8262.3            | 8266.2           | 8253.1            |
| BIC                | 8363.6           | 8359.8            | 8370.3            | 8366.7            | 8366.4            | 8370.3           | 8376.2            |

GAS parameter estimates and goodness-of-fit statistics for several SMXVAR model specifications. Standard errors are given in parentheses.

### 5.2.1 SMVAR Model Results

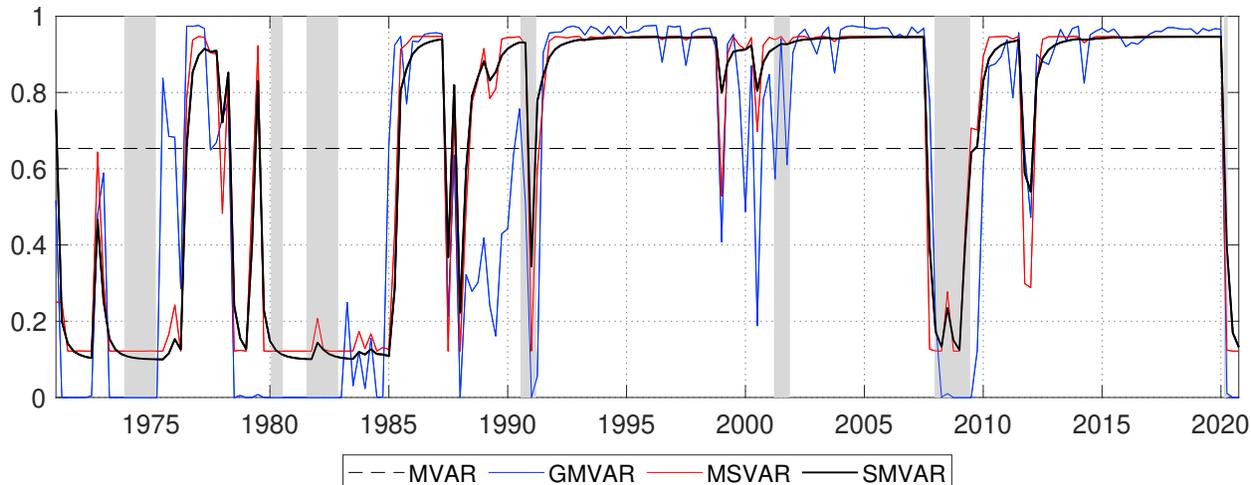
We choose the quarter-over-quarter annualized growth rate of real GDP as indicator for real economic activity<sup>17</sup> and we rely on the National Financial Conditions Index (NFCI) provided by the Federal Reserve Bank of Chicago to capture financial conditions.<sup>18</sup> Our sample starts in 1971q1 with the first observation of the NFCI and ends in 2020q4. NBER recessions are accompanied by lower GDP growth and a tightening of financial conditions as indicated by

<sup>17</sup>Source: <https://fred.stlouisfed.org/series/GDPC1>.

<sup>18</sup>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. The data set and some background can be found here: <https://www.chicagofed.org/publications/nfci/index>.

higher values of the NFCI with the mild and short recession of 2001 being an exception.<sup>19</sup> In addition, the Covid-19 recession is also only associated with a short-lived surge in the NFCI.

**Figure 6:** Development of  $\alpha_t$  over Time for Macro-Financial Linkages



Probability of being in the normal state as indicated by the estimated mixture weights  $\alpha_t$  based on four different approaches. Shaded areas are NBER recessions.

Again, we estimate a two-state SMVAR with one lag to examine the joint distribution of the NFCI and real GDP as suggested by both information criteria. The fitted transformed mixture weights  $\alpha_t$  are shown in Figure 6. All recessions (except the one in 2001) are characterized by a low value of the mixtures weights  $\alpha_t$ . In contrast to the stock-bond application, the mixture weight varies quite clearly between two levels. This leads – similar to MSVAR models – to a switching interpretation of the two regimes as opposed to a mixture of two components. 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.

<sup>19</sup>Figure C.2 shows both series over time.

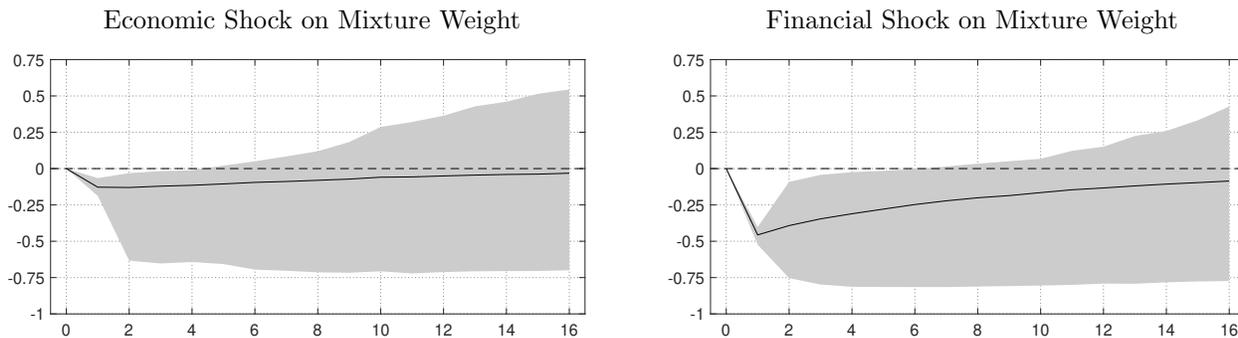
**Table 8:** Estimation Results for Macro-Financial Linkages

| <b>(a) VAR Parameters</b>           |                    |                  |                   |                   |
|-------------------------------------|--------------------|------------------|-------------------|-------------------|
|                                     | Normal Regime      |                  | Crisis Regime     |                   |
|                                     | $NFCI_t$           | $\Delta GDP_t$   | $NFCI_t$          | $\Delta GDP_t$    |
| $NFCI_{t-1}$                        | 0.749<br>(0.027)   | 0.016<br>(0.342) | 0.751<br>(0.086)  | 0.800<br>(1.589)  |
| $\Delta GDP_{t-1}$                  | -0.013<br>(0.003)  | 0.106<br>(0.052) | -0.021<br>(0.008) | -0.206<br>(0.149) |
| <i>const</i>                        | -0.143<br>(0.016)  | 0.020<br>(0.002) | 0.303<br>(0.137)  | 0.051<br>(0.013)  |
| <b>(b) (Co-)Variance Parameters</b> |                    |                  |                   |                   |
|                                     | Normal Regime      |                  | Crisis Regime     |                   |
|                                     |                    |                  |                   |                   |
| $\Omega_{1,1}$                      | 0.014<br>(0.002)   |                  | 0.585<br>(0.108)  |                   |
| $\Omega_{1,2}$                      | 0.0003<br>(0.0002) |                  | -0.008<br>(0.007) |                   |
| $\Omega_{2,2}$                      | 0.0003<br>(0.0000) |                  | 0.005<br>(0.001)  |                   |
| <b>(c) GAS Parameters</b>           |                    |                  |                   |                   |
|                                     | $\bar{\alpha}$     | $a$              | $b$               |                   |
|                                     | 1.123<br>(0.811)   | 3.327<br>(0.781) | 0.901<br>(0.046)  |                   |

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 given in parentheses.

Table 8 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 shown 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 8 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 therefore interpret the latter as crisis regime. The NFCI is similarly persistent in both 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

**Figure 7:** Impulse Responses of Mixture Weight



Generalized impulse responses (black lines) of the mixture weight to an economic and a financial shock of four SD starting in the normal regime. Grey-shaded areas indicate 68% confidence bands.

to be persistent in the normal state, whereas there is no significant persistence during crisis times. Turning to the GAS parameters, we find a high degree of persistence in the weights as indicated by the estimate of 0.901 for the autoregressive coefficient  $b$  of Eq. (20). This is also the reason for the smooth development of  $\alpha_t$  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 is 0.755. Consequently, the economy is – on average – in 75% of the time in the normal state.

### 5.2.2 Impulse Responses

Given the results of Application 2, a model with Markov-switching weights would also be able to represent the reported dynamics. However, the SMVAR has the unique feature that weights are driven by previous observations. This allows to investigate the impact of a shock in an impulse response analysis not only via the direct lags in the VAR components but also by the lagged impact on the component weights. In particular, we can investigate the impact of an economic or financial shock on the probability of a switch into the crisis regime. Figure 7 shows the GIRFs to an economic and a financial shock of four SD starting in the normal regime (based on 10,000 Monte Carlo replications). We notice that the likelihood

of triggering a crisis is – on average – higher after an financial shock than after a economic shock of equal relative size. More precisely, the probability of a crisis increases by roughly 45 percentage points (pp) one period after a shock to financial conditions as compared to 13 pp in the case of an economic growth shock. Given an initial normal times regime weight of 90%, the probability of switching into the crisis increases from 10% to 55% (23 %) in the case of a financial (economic) shock.

We provide a detailed impulse response analysis in Appendix C.5. In particular, the results show that the impulse responses become more uncertain in dynamic mixture VAR models – such as the SMVAR – when compared to a situation where regime switches have been ruled out. This is reasonable since the trajectory of the impulse response crucially depends on whether or not the shock triggers a transition into another regime.

## 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 GMVAR always updates the weights according to the recent scaled observation density. In addition, the SM(X)VAR is also helpful if the instruments driving the mixture weights in a Logit MVAR are unknown or uncertain as it outperforms the Logit MVAR if the instruments are too noisy.

The benefits of the SM(X)VAR model were illustrated using two empirical applications with time series of different frequency and noise level. The model is particularly appealing because of its high flexibility. Moreover, it also provides novel analytical tools, such as measuring the impact of shocks on mixture weights as part of an impulse response analysis and incorporating predictors in the weight updating process.

## 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., 2023. The Financial Accelerator in the Euro Area: New Evidence Using a Mixture VAR Model. *Macroeconomic Dynamics* 27, 1893–1931.
- 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., Francq, C., Laurent, S., 2023. Quasi score-driven models. *Journal of Econometrics* 234, 251–275.
- 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.
- Byrd, R.H., Hribar, M.E., Nocedal, J., 1999. An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization* 9, 877–900.
- 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.
- Guidolin, M., Ono, S., 2006. Are the dynamic linkages between the macroeconomy and asset prices time-varying? *Journal of Economics and Business* 58, 480–518.
- Guidolin, M., Timmermann, A., 2006. An econometric model of nonlinear dynamics in the joint distribution of stock and bond returns. *Journal of Applied Econometrics* 21, 1–22.
- Haas, M., Pigorsch, C., 2009. Financial economics, fat-tailed distributions. *Encyclopedia of Complexity and Systems Science* 4, 3404–3435.
- 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.
- Kole, E., van Dijk, D., 2023. Moments, shocks and spillovers in Markov-switching VAR models. *Journal of Econometrics* 236, 105474.

- 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 and Heidelberg.
- Kullback, S., Leibler, R.A., 1951. On information and sufficiency. *Annals of Mathematical Statistics* 22, 79–86.
- McLachlan, G., Peel, D., 2004. *Finite Mixture Models*. Wiley Series in Probability and Statistics, Wiley, Hoboken.
- 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.

# Online Appendix

## A Derivations for GAS Updating Equations

### A.1 General Weight Updating

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

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

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

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

$$= \mathcal{J}_h(\tilde{\alpha}_t) \cdot \begin{pmatrix} \frac{f_1(y_t | \mathcal{F}_{t-1})}{f(y_t | \mathcal{F}_{t-1})} \\ \vdots \\ \frac{f_K(y_t | \mathcal{F}_{t-1})}{f(y_t | \mathcal{F}_{t-1})} \end{pmatrix} \quad (\text{A.5})$$

### A.2 Jacobian of Logit Transformation

$k = l, k \neq K$ :

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

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

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

$k \neq l, k \neq K$ :

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

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

$$(\text{A.11})$$

$k = K$ :

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

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

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

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

## B Additional Simulation Results

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

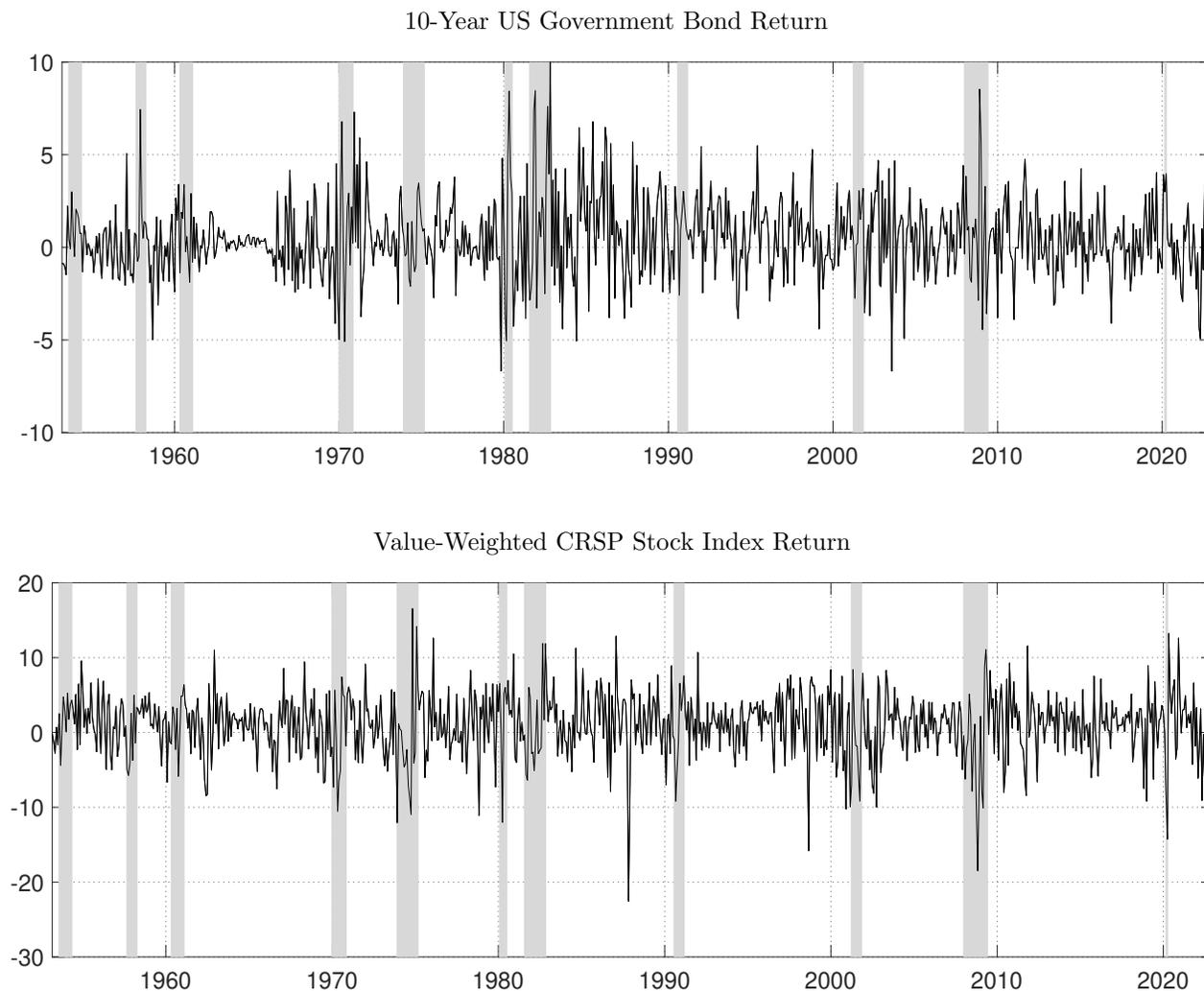
| DGP        | $T$  | $L$     |         |         |         | AIC    |        |        |        | BIC    |        |        |        |
|------------|------|---------|---------|---------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
|            |      | MVAR    | MSVAR   | GMVAR   | SMVAR   | MVAR   | MSVAR  | GMVAR  | SMVAR  | MVAR   | MSVAR  | GMVAR  | SMVAR  |
| I Constant | 200  | -430.9  | -428.2  | -450.2  | -430.2  | 899.9  | 898.5  | 938.3  | 902.4  | 962.5  | 967.7  | 1001.0 | 971.6  |
|            | 500  | -1093.2 | -1090.6 | -1144.0 | -1093.1 | 2224.3 | 2223.1 | 2325.9 | 2228.3 | 2304.4 | 2311.6 | 2406.0 | 2316.8 |
|            | 1000 | -2194.2 | -2191.4 | -2297.9 | -2195.3 | 4426.3 | 4424.9 | 4633.9 | 4432.6 | 4519.6 | 4527.9 | 4727.1 | 4535.7 |
| II Break   | 200  | -433.2  | -422.8  | -435.5  | -421.5  | 904.5  | 887.5  | 909.0  | 885.0  | 967.13 | 956.8  | 971.7  | 954.3  |
|            | 500  | -1091.7 | -1077.1 | -1101.8 | -1058.9 | 2221.3 | 2196.3 | 2241.6 | 2159.8 | 2301.4 | 2284.8 | 2321.7 | 2248.3 |
|            | 1000 | -2188.0 | -2167.3 | -2211.1 | -2110.2 | 4414.0 | 4376.6 | 4460.1 | 4262.3 | 4507.2 | 4479.6 | 4553.4 | 4365.4 |
| III Cycle  | 200  | -433.6  | -419.9  | -432.9  | -418.8  | 905.2  | 881.7  | 903.8  | 879.6  | 967.9  | 951.0  | 966.4  | 948.8  |
|            | 500  | -1090.4 | -1070.0 | -1092.5 | -1045.5 | 2218.8 | 2182.0 | 2223.0 | 2133.0 | 2298.8 | 2270.5 | 2303.1 | 2221.5 |
|            | 1000 | -2184.0 | -2153.3 | -2191.3 | -2088.6 | 4406.1 | 4348.6 | 4420.7 | 4219.2 | 4499.3 | 4451.7 | 4513.9 | 4322.2 |
| IV MS      | 200  | -373.4  | -346.3  | -362.9  | -351.7  | 784.8  | 734.6  | 763.8  | 745.3  | 847.51 | 803.8  | 826.4  | 814.6  |
|            | 500  | -961.0  | -889.5  | -925.8  | -896.5  | 1959.9 | 1821.0 | 1889.5 | 1834.9 | 2040.0 | 1909.5 | 1969.6 | 1923.4 |
|            | 1000 | -1931.6 | -1789.0 | -1857.1 | -1798.6 | 3901.2 | 3620.0 | 3752.1 | 3639.2 | 3994.5 | 3723.1 | 3845.4 | 3742.3 |
| V LAR      | 200  | -421.9  | -415.5  | -430.9  | -414.3  | 881.7  | 873.1  | 899.9  | 870.7  | 944.4  | 942.3  | 962.5  | 939.9  |
|            | 500  | -1083.3 | -1070.7 | -1100.5 | -1057.2 | 2204.5 | 2183.5 | 2239.1 | 2156.3 | 2284.6 | 2272.0 | 2319.1 | 2244.8 |
|            | 1000 | -2182.4 | -2158.6 | -2212.0 | -2123.1 | 4402.7 | 4359.2 | 4461.9 | 4288.2 | 4495.9 | 4462.2 | 4555.2 | 4391.3 |

Average log-likelihood ( $L$ ), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) for the estimate of the mixture weight  $\alpha_{1,t}$  from MVAR, MSVAR, GMVAR, and SMVAR models with 1000 Monte Carlo replications and different sample sizes  $T$ .

# C Additional Application Results

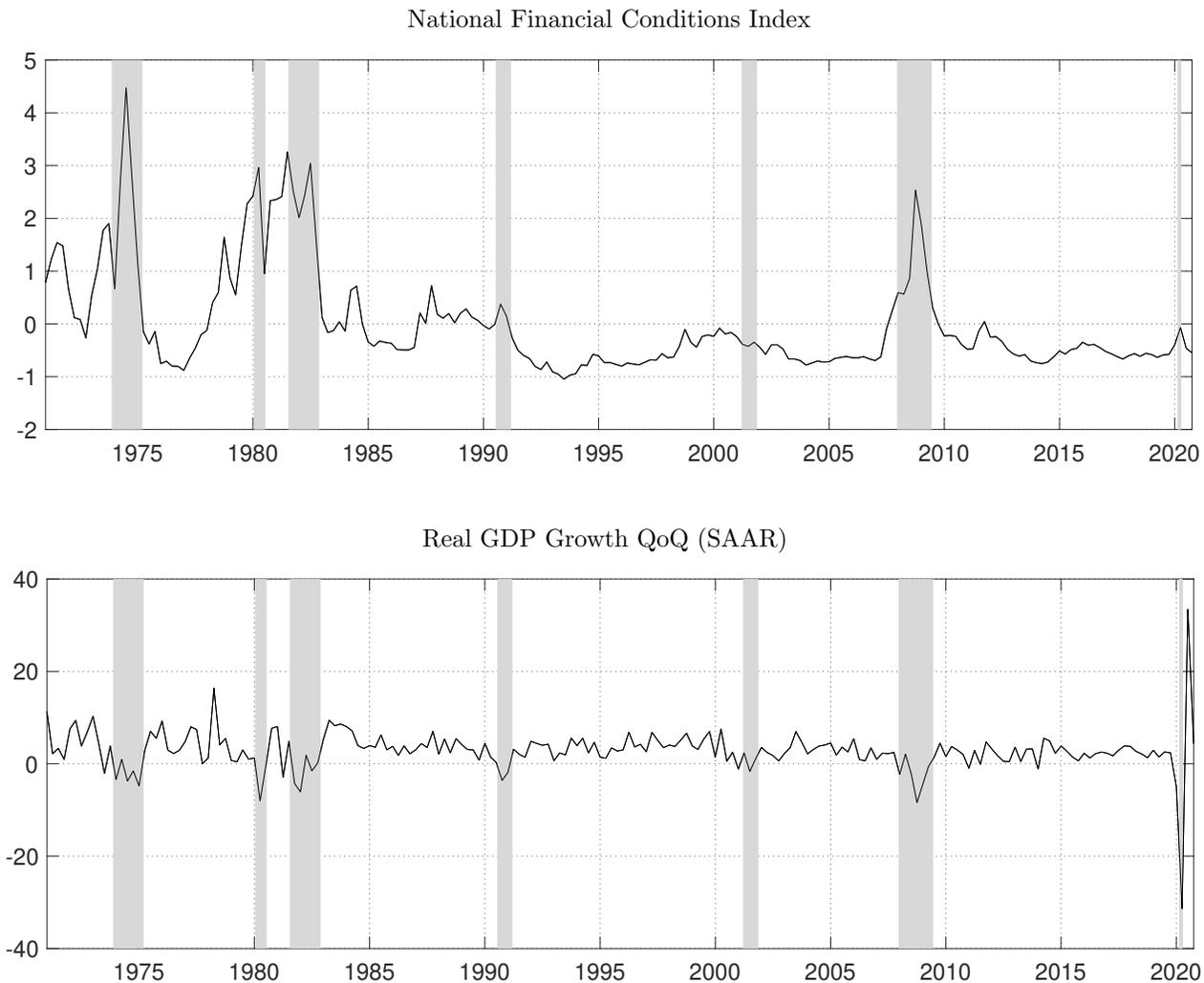
## C.1 Time Series Plots

Figure C.1: Application 1: Excess Return Series



The upper panel shows a bond index return based on 10-year to maturity US government bonds. The lower panel shows the return to a value-weighted stock index (including dividends) from the NYSE, NASDAQ, and AMXE exchanges (provided by CRSP). Returns are in excess of the 1-month T-Bill rate. Shaded areas indicate NBER recessions.

**Figure C.2:** Application 2: NFCI and Real GDP Growth



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 indicate NBER recessions.

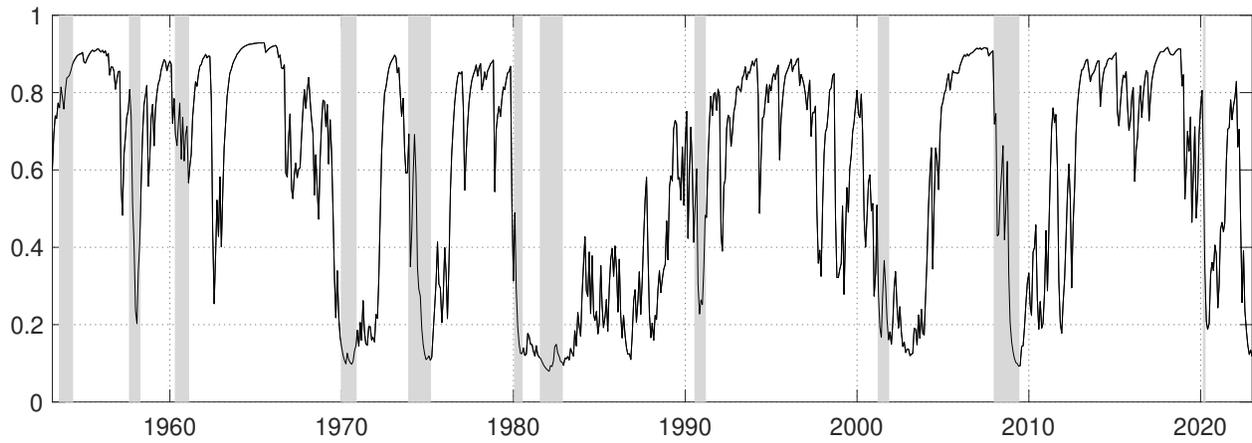
## C.2 Application 1: tSMVAR Results

**Table C.1:** Estimation Results for Stock and Bond Return Dynamics (tSMVAR)

| <b>(a) VAR Parameters</b>            |                    |                   |                   |                   |
|--------------------------------------|--------------------|-------------------|-------------------|-------------------|
|                                      | Normal Regime      |                   | Crisis Regime     |                   |
|                                      | $Bond_t$           | $Stock_t$         | $Bond_t$          | $Stock_t$         |
| $Bond_{t-1}$                         | 0.067<br>(0.049)   | -0.049<br>(0.023) | 0.087<br>(0.062)  | -0.083<br>(0.031) |
| $Stock_{t-1}$                        | 0.194<br>(0.116)   | -0.067<br>(0.057) | 0.201<br>(0.137)  | 0.100<br>(0.069)  |
| $const$                              | 0.074<br>(0.076)   | 1.289<br>(0.198)  | 0.234<br>(0.166)  | -0.291<br>(0.371) |
| <b>(b) (Co-)Variance Parameters</b>  |                    |                   |                   |                   |
|                                      | Normal Regime      |                   | Crisis Regime     |                   |
| $\Omega_{1,1}$                       | 1.684<br>(0.206)   |                   | 7.224<br>(0.963)  |                   |
| $\Omega_{1,2}$                       | -0.006<br>(0.275)  |                   | 2.131<br>(0.914)  |                   |
| $\Omega_{2,2}$                       | 7.792<br>(0.969)   |                   | 29.039<br>(3.817) |                   |
| <b>(c) GAS Parameters</b>            |                    |                   |                   |                   |
|                                      | $\bar{\alpha}$     | $a$               | $b$               |                   |
|                                      | 0.407<br>(0.501)   | 2.076<br>(0.632)  | 0.953<br>(0.018)  |                   |
| <b>(d) Distributional Parameters</b> |                    |                   |                   |                   |
|                                      | $\nu$              |                   |                   |                   |
|                                      | 22.854<br>(16.602) |                   |                   |                   |

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 given in parentheses.

**Figure C.3:** Development of  $\alpha_t$  over Time for Stock and Bond Return Dynamics (tSMVAR)



Probability of being in the normal state as indicated by the estimated mixture weights  $\alpha_t$ . Shaded areas indicate NBER recessions.

### C.3 Application 2: Detailed Impulse Response Analysis

In the following, we discuss the impulse response analysis for the application on macro-financial linkages in more detail. Algorithm 1 shows the steps to simulate the GIRF given by Eq. (17).

---

**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**  $r = 1, \dots, R$  **do**
  - 2.a) Draw  $N + 1$  random shocks  $\varepsilon_t^{(r)}, \varepsilon_{t+1}^{(r)}, \dots, \varepsilon_{t+N}^{(r)}$  from  $\mathcal{N}(0, I_d)$ .
  - 2.b) **for**  $n = 0, \dots, N$  **do**
    - 2.b) i) Use  $\varepsilon_{t+n}^{(r)}$  to compute  $(y_{t+n}^{(r)}(\mathcal{F}_{t-1}))$  with equation (1)
    - 2.b) ii) Iterate the SMVAR recursion (3) to compute  $\alpha_{t+n+1}^{(r)}$
    - 2.b) iii) Use  $\alpha_{t+n+1}^{(r)}$  to draw  $s_{t+n+1}^{(r)}$ .
  - end**
  - 2.c) Redo 2.b) with  $\varepsilon_t^{(r)} = \xi_t$  and shocks  $(\varepsilon_{t+n}^{(r)})_{n=1}^N$  to compute  $(y_{t+n}^{(r)}(\xi_t, \mathcal{F}_{t-1}))_{n=0}^N$ .
  - 2.d) Calculate
 
$$GI^{(r)}(n, \xi_t, \mathcal{F}_t) = y_{t+n}^{(r)}(\xi_t, \mathcal{F}_{t-1}) - y_{t+n}^{(r)}(\mathcal{F}_{t-1})$$
 for  $n = 0, \dots, N$ .
- end**
- 3.) Form averages over the Monte Carlo replications:

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


---

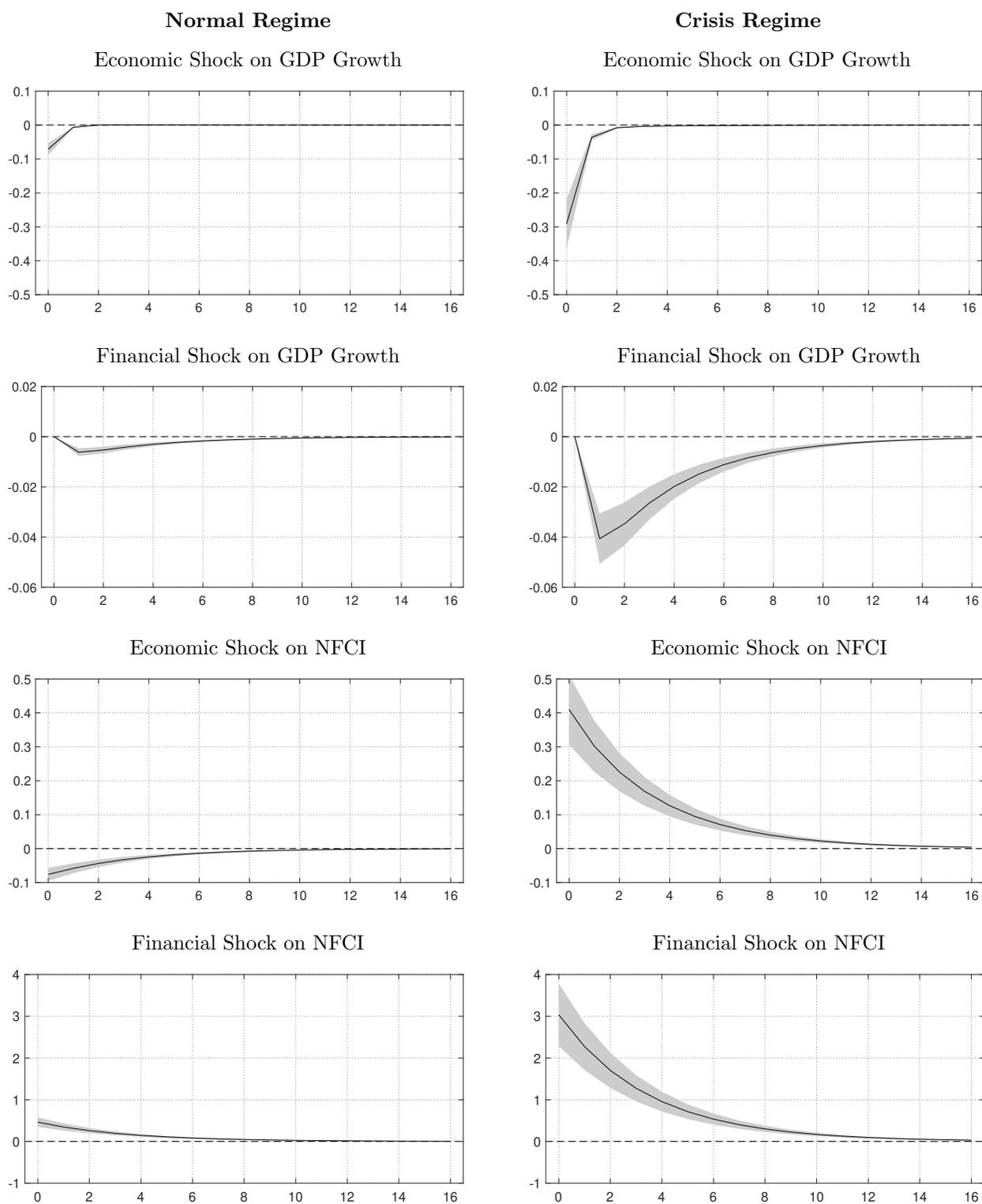
To avoid conditioning on an information set, another Monte Carlo loop might be added around Algorithm 1 in which  $\widehat{GI}$  is integrated over draws from the historical sequence  $(y_i)_{i=0}^{t-1}$ . Moreover, the algorithm could be easily adjusted for different distributions that are possible in the SMVAR framework by replacing the innovation distribution in step 2.a).

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 C.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 persistence as its impact dies out only slowly. The financial shock also reduces GDP growth significantly with a peak effect of  $-0.6$  basis points (bps).

The right panel of Figure C.4 shows the GIRFs that occur within the crisis 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 ( $-4.1$  bps as compared to  $-0.6$  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 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  $0.410$  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).

**Figure C.4:** Component-wise Impulse Responses for Macro-Financial Linkages

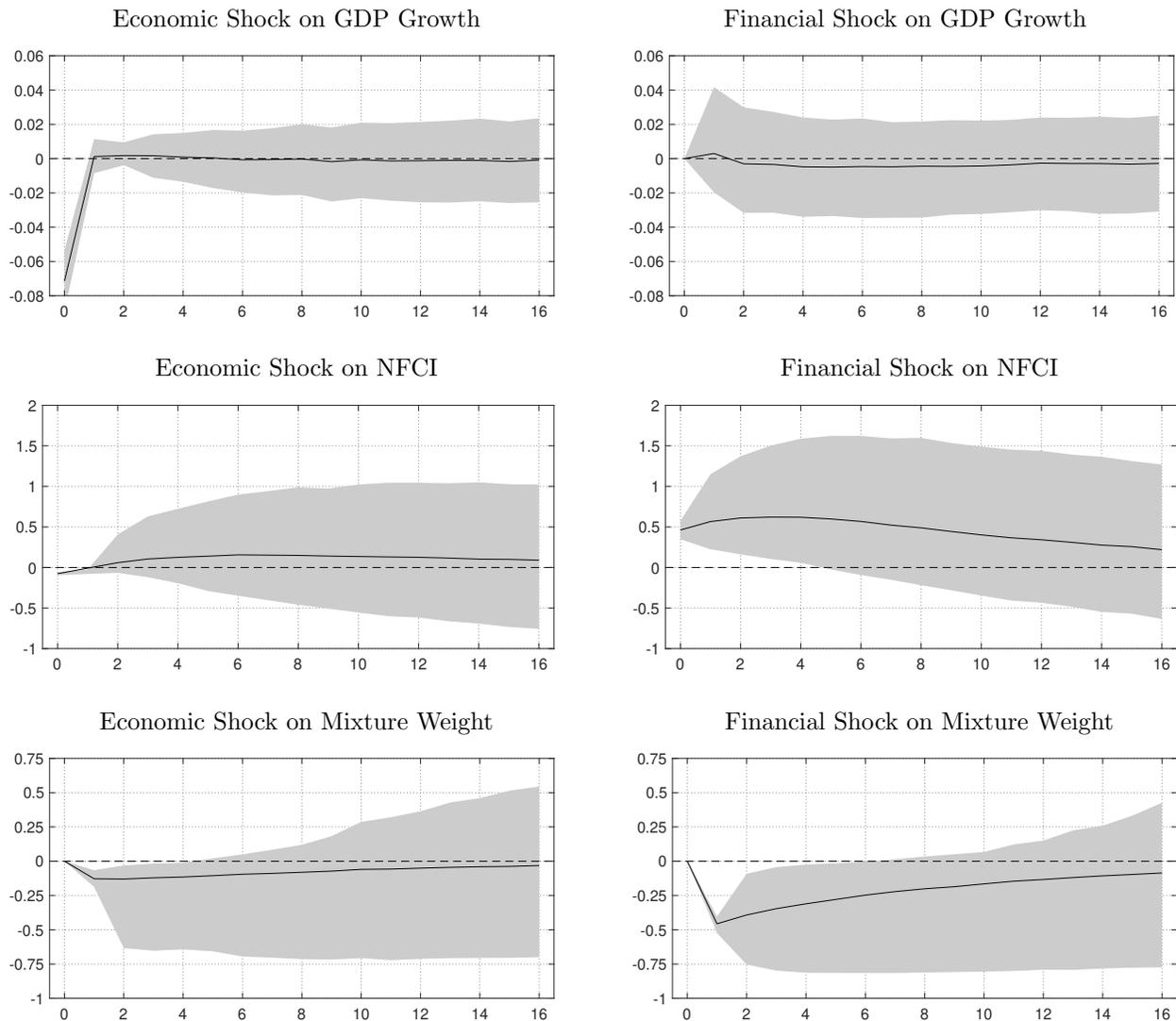


Generalized impulse responses (black lines) to an economic and a financial shock of four SD within the normal regime (left panel) and the crisis regime (right panel). Regime shifts are ruled out. Grey-shaded areas indicate 68% confidence bands.

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 (cf. Figure 7 and the discussion in Section 5.2.2). Figure C.5 shows the full set of GIRFs to an economic and a financial shock of four SD starting in the normal regime.

There are four striking findings. First, the qualitative results concerning the persistence of the shocks are similar when introducing possible regime shifts. The economic shock decreases the shock variable significantly on impact and the effect is rapidly dying out (after 1 quarter), whereas the financial shock decreases the shock variable significantly on impact with a clearly slower decay (4 quarters). 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 C.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. This is particularly prevalent when analyzing the effect of a financial shock on GDP growth. In contrast to Figure C.4, the response is insignificant over the full horizon.

**Figure C.5:** Impulse Responses for Macro-Financial Linkages with Mixture Dynamics



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. The bottom panel replicates Figure 7 from Section 5.2.2.

Third, we also notice that the likelihood of triggering a crisis is – on average – higher after an financial shock than after a economic shock of equal relative size. More precisely, the probability of a crisis increases by roughly 45 percentage points (pp) one period after a shock to financial conditions as compared to 13 pp in the case of an economic growth shock. Given an initial normal times regime weight of 90%, the probability of switching into the

crisis increases from 10% to 55% (23 %) in the case of a financial (economic) 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 slightly positive in the aftermath. This is in contrast to the crisis regime in which the response is largely positive on impact and decays only slowly.

These observed dynamics are 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. For instance, an initial shock transmits directly through the normal times VAR model connections, but additionally increases the probability of shifting into a crisis. Hence, the subsequent trajectory will also be affected by a change in the unconditional mean and variance if the higher crisis probability triggers an actual shift into the crisis regime.