Assessing Asymmetric Macroeconomic Risk

Stéphane Lhuissier*

July 2, 2025

Abstract

I propose a dynamic factor model with time-varying skewness to assess asymmetric risk around the economic outlook across a set of macroeconomic aggregates. Applied to U.S. data, the model shows that macroeconomic skewness is procyclical, displays significant independent variations from GDP growth skewness, and does not require conditioning on financial variables to manifest. Compared to univariate benchmarks, the model improves the detection of downside risk to growth and delivers more accurate predictive distributions, especially during downturns. These findings underscore the value of using a richer information set to quantify the balance of macroeconomic risks.

Keywords: Dynamic Factor Models, Markov-switching, Skewness.

JEL Codes: C34, C38, C53, E37.

*Banque de France, 31, Rue Croix des Petits Champs, DGSEI-DEMFI-POMONE 41-1422, 75049 Paris Cedex 01, FRANCE (Email: stephane.lhuissier@hotmail.com; URL: http://www.stephanelhuissier.eu). The views expressed in this paper are those of the author and should under no circumstances be interpreted as reflecting those of the Banque de France or the Eurosystem. I owe a special thanks to Todd Clark and Pablo Guerrón-Quintana (discussant) for extensive comments on an earlier draft of the paper. I also thank Borağan Aruoba, Daniel Lewis, Jean-Paul Renne, and seminar participants at the Banque de France, for helpful comments.

1 Introduction

In recent years, central banks have increasingly adopted a risk management perspective in the formulation of monetary policy. Rather than considering solely the most likely macroeconomic outcome, policymakers often take into account the full distribution of possible outcomes, placing particular attention on tail risks. Central banks discuss whether they see risks around the outlook as balanced, on the upside or on the downside.

This shift in perspective has spurred the development of empirical tools designed to quantify asymmetric risks to real gross domestic product (GDP) growth. A prominent example is the growth-at-risk (GaR) framework developed by Adrian, Boyarchenko, and Giannone (2019), which uses quantile regressions to show that risk to real GDP growth becomes more left-skewed when financial conditions tighten. Related work has employed alternative methods, such as stochastic volatility or Markov-switching models.[1]

Although the literature has made important progress in documenting time-varying asymmetric risks to real GDP growth, much less attention has been paid to asymmetric risks for the broader economy. GDP growth, while central to understanding the business cycle, reflects only a single dimension of economic activity. It does not capture the broader array of economic indicators — such as output, employment, and sales — that summarize the macroeconomy. Consequently, using only GDP growth to assess the balance of macroeconomic risks can obscure asymmetries in other critical areas that also reflect broader economic risks.

Against this background, I propose and illustrate a framework for the assessment of broad-based macroeconomic risks in a systematic and replicable manner. My framework

[1] A growing and non-exhaustive list of papers includes Giglio, Kelly, and Pruitt (2016), Plagborg-Møller, Reichlin, Ricco, and Hasenzagl (2020), Adrian, Grinberg, Liang, Malik, and Yu (2022), Caldara, Cascaldi-Garcia, Cuba-Borda, and Loria (2022), Lhuissier (2022), Wolf (2023), Caldara, Scotti, and Zhong (2024), Carriero, Clark, and Marcellino (2024), Delle Monache, De Polis, and Petrella (2024), Forni, Gambetti, and Sala (2024), Loria, Matthes, and Zhang (2024) and Castelnuovo and Mori (2025).

More broadly, there is an increasing literature dealing with fluctuations in asymmetric macroeconomic shocks. Notable examples include Rietz (1988), Barro (2009), Barro and Ursúa (2012), Gabaix (2012), Gourio (2012), Busch, Domeij, Guvenen, and Madera (2022), Cai and Guerrón-Quintana (2023), Salgado, Guvenen, and Bloom (2023) and Bekaert, Engstrom, and Ermolov (forthcoming).

has three key ingredients. First, I work with a dynamic factor model (DFM) that treats economic activity as a latent factor, inferred from a set of macroeconomic indicators. This approach aligns with the theoretical perspective (e.g., Lucas, 1977) that the business cycle is not captured by any single observable series but rather reflects the shared dynamics underlying many indicators. By extracting a common factor, the model provides a coherent representation of the aggregate economy. Second, I assume an asymmetric distribution for the factor disturbances. Specifically, I employ the closed skew-normal distribution developed by González-Farías, Domínguez-Molina, and Gupta (2004), which allows for departure from symmetry. Beyond introducing skewness, this distribution remains closed under marginalization and conditioning, thereby preserving key tractability features. Third, I allow for time variation in the parameters of the factor distribution. Time variation is modeled as a Markov-switching process, following Hamilton (1989)'s work. The inclusion of a Markov switching structure is particularly valuable in this context, as it allows the model to capture abrupt changes in the behavior of macroeconomic data, such as those typically observed around recessions.

Together, these components allow for capturing business-cycle variation in the conditional distribution and higher order moments of the common macroeconomic factor. In particular, a measure, or index, of *macroeconomic skewness* can be constructed by deriving analytically the conditional skewness of the macroeconomic factor, which reflects variations in the balance of risks of a set of macroeconomic aggregates. Although the model is neither linear nor Gaussian, it can be estimated by a modified Kalman filter that I develop in the paper. Using synthetic data, I demonstrate the effectiveness of the filtering and estimation procedure in the online Appendix.

I apply my general framework to U.S. data. I use a set of four macroeconomic aggregates (i.e., real GDP growth, real personal income, real manufacturing and trade sales, and employment), which are the same four sectoral variables typically utilized in this literature (e.g., Stock and Watson, 1991; Chauvet, 1998). My evidence shows substantial cyclical variation

in the broad-based balance of risks: macroeconomic skewness displays a procyclical pattern, with a tendency to rapidly decline to negative territory during downturns and to rise during expansions. Skewness tends also to decline in anticipations of recessions.

Although my measure shares some commonalities with skewness computed solely from GDP growth, it also exhibits substantial differences, indicating that GDP-based skewness does not fully capture economy-wide asymmetric risks. Furthermore, the measure is partially correlated with GDP growth skewness that conditions on past economic and financial conditions (e.g., Adrian, Boyarchenko, and Giannone, 2019). This is important because it suggests that financial conditions are not necessarily a driver of asymmetric macroeconomic risks.

To evaluate the practical relevance of the skewness measure produced by the model, I assess the model's out-of-sample performance in predicting risks to GDP growth. The focus lies on whether incorporating a broad set of macroeconomic indicators via my framework improves predictive accuracy relative to univariate benchmark approaches: a univariate version of the model and the quantile regression framework of Adrian, Boyarchenko, and Giannone (2019). My framework delivers sharper signals of downside risk around U.S. recessions, particularly the 2001 and 2008 episodes, when univariate benchmark models either mis-measure downside risks or understate recession probabilities. These improvements reflect the value of leveraging information from multiple macroeconomic indicators beyond GDP alone.

A formal forecast evaluation confirms the model's relative gains. While full-sample improvements in density and quantile forecast metrics are modest and statistically insignificant, the model clearly outperforms benchmarks during recession periods. Log scores improve significantly over univariate benchmarks, and the model demonstrates better accuracy in predicting the lower tail of GDP growth. Moreover, calibration analysis confirms the reliability of the forecasts, showing that they generally align well with actual economic outcomes. These findings underscore the usefulness of the proposed framework for tracking and quantifying macroeconomic risks.

Relation to other studies. My work adds to the limited body of research dedicated to the assessment of broad-based macroeconomic risks. Caldara, Mumtaz, and Zhong (2024) uses a DFM with stochastic volatility to quantify risk around the outlook using a broad dataset of macroeconomic and financial indicators. The authors provide evidence of asymmetric risk dynamics in the common factors, in the sense that downside risk varies more over time than upside risk. This arises from conditional predictive distributions that are symmetric and subject to simultaneous changes in mean and variance. My results are complementary to this paper as I provide evidence of time variation in the macroeconomic factor skewness, and thus in the broad-based balance of risks. Most closely related to my work is the paper by Iseringhausen, Petrella, and Theodoridis (forthcoming), which measures broad-based macroeconomic skewness using principal component analysis applied to a set of variable-specific skewness measures. Compared to my method, this approach has the advantage of using a large dataset containing 210 time series, thus providing a measure of macroeconomic skewness for the broader economy. The strength of my method, instead, lies in proposing a model-based alternative that leverages the structure of the transition equation for the estimation of the macroeconomic factor and its time-varying skewness.

From a methodological standpoint, this paper is related to an existing literature on DFMs, which have been instrumental in capturing commonalities across sectors or economies (e.g., Stock and Watson, 1991; Kose, Otrok, and Whiteman, 2003; Ciccarelli and Mojon, 2010). A number of studies has extended the standard DFM approach by allowing time-varying parameters. Building upon ideas in Kim (1993), Kim and Yoo (1995) and Chauvet (1998) propose a Markov-switching DFM of coincident economic indicators. As an alternative to Markov-switching models, the literature has also considered model parameters evolving according to a multivariate autoregressive process. For example, Del Negro and Otrok (2008) and Mumtaz and Surico (2012) estimate DFMs with time-varying parameters and stochastic volatility. Antolin-Diaz, Drechsel, and Petrella (2017) propose a DFM including time-varying long-run growth and stochastic volatility, and Antolin-Diaz, Drechsel, and Petrella (2024)

further introduce Student-t distributed outliers. Guerrón-Quintana, Khazanov, and Zhong (2024) estimate DFMs that allow for general nonlinearities in the transition and measurement equations. However, none of these papers explicitly consider time variation beyond the first two moments. To the best of my knowledge, my paper is the first to incorporate time-varying skewness within a DFM framework.

The plan of the paper is as follows. Section 2 provides a detailed statement of the dynamic factor framework and in Section 3 I represent it as a state-space filtering problem. Section 4 discusses the data, model implementation, and adjustments to incorporate the Covid-19 period. Section 5 presents the estimates of the model and Section 6 reports the resulting macroeconomic skewness and discusses how it differs from alternative skewness measures. Section 7 reports the results of out-of-sample forecasting exercises. Section 8 concludes the paper.

2 The Modeling Framework

Here, I propose a DFM that allows the factor distribution to be time-varying and asymmetric. While the model itself is conceptually straightforward, its features introduce significant complexity in the extraction of the latent factor and in the estimation of parameters of the model — an issue addressed in detail in the next section of the paper.

Let y_{it} denote the ith macroeconomic indicator at date t, which consists of an individual component z_{it} and a linear combination of current and lagged values of a common macroeconomic factor n_t:

$$y_{it} = \gamma_i(L)n_t + z_{it}, \tag{1}$$

with $\gamma_i(L)$ is a scalar lag polynomial, L denotes the lag operator, $i = 1, \ldots, N$, and $t = 1, \ldots, T$.

The evolution of the idiosyncratic term z_{it} can be described by a Gaussian autoregressive

(AR) representation:

$$\psi_i(L) z_{it} = \varsigma_{it}, \tag{2}$$

$$\varsigma_{it} \sim \text{normal}\left(0, \sigma^2_{i,\varsigma}\right), \tag{3}$$

where normal (a, b) refers to the normal probability density function (pdf) with mean a and covariance matrix b.

The common factor is assumed to be generated by a skewed AR process with Markov-switching:

$$\phi(L) n_t = \vartheta_t, \tag{4}$$

$$\vartheta_t \sim \text{closed skew-normal}(\mu_{s_t^{\text{location}},\vartheta}, \sigma^2_{s_t^{\text{scale}},\vartheta}, \alpha_{s_t^{\text{shape}},\vartheta}, 0, 1), \tag{5}$$

where closed skew-normal $(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Gamma}, \boldsymbol{\nu}, \boldsymbol{\Delta})$ refers to the closed skew-normal distribution introduced by González-Farías, Domínguez-Molina, and Gupta (2004), with $\boldsymbol{\mu}$ the location parameter, $\boldsymbol{\Sigma}$ the scale parameter, $\boldsymbol{\Gamma}$ the shape parameter governing the skewness, and $\boldsymbol{\nu}$ and $\boldsymbol{\Delta}$ two parameters which are opens to interpretation. This distribution allows for departure from symmetry and has the advantage of being closed under marginalization, conditional distributions, linear transformation, sums of independent random variables from this family, and joint distribution of independent random variables in this family.[2]

[2]Compared to the normal distribution, the closed skew-normal is a distribution that has several additional parameters: a shape parameter $\boldsymbol{\Gamma} \in \mathbb{R}$, which allows for possible deviation from symmetry, and two other parameters $\boldsymbol{\nu} \in \mathbb{R}$ and $\boldsymbol{\Delta} \in \mathbb{R}$, which are opens to interpretation. The two latter parameters are essential because they allow for closure of the distribution under conditioning and marginalization. A random vector $\boldsymbol{Y}$ distributed according to the multivariate closed skew-normal with parameters: $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\nu}$, $\boldsymbol{\Delta}$, denoted by $\boldsymbol{Y} \sim$ closed skew-normal $(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Gamma}, \boldsymbol{\nu}, \boldsymbol{\Delta})$ has the probability density function given by:

$$p(\boldsymbol{Y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Gamma}, \boldsymbol{\nu}, \boldsymbol{\Delta}) = \frac{\boldsymbol{\Phi}\left(\boldsymbol{\Gamma}(\boldsymbol{Y} - \boldsymbol{\mu}); \boldsymbol{\nu}, \boldsymbol{\Delta}\right)}{\boldsymbol{\Phi}\left(\mathbf{0}; \boldsymbol{\nu}, \boldsymbol{\Delta} + \boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\Gamma}'\right)} \boldsymbol{\phi}(\boldsymbol{Y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

where $\boldsymbol{\Phi}(.; \boldsymbol{m}, \boldsymbol{S})$ is the cumulative density function of the multivariate normal distribution with expectation vector $\boldsymbol{m}$ and covariance matrix $\boldsymbol{S}$; and $\boldsymbol{\phi}(.; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the probability density function of a multivariate normal distribution with expectation vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. If the shape parameter is equal to zero, then the density of $\boldsymbol{Y}$ is a multivariate normal distribution with mean $\boldsymbol{\mu}$, and variance $\boldsymbol{\Sigma}$. Another special case is given by $y \sim$ closed skew-normal $(0, 1, \gamma, 0, 1)$, which corresponds to the univariate standardized

Each parameter of the factor distribution (location, scale, and shape) is dependent upon an unobserved variable s_t^{h}, with $h \in \{\text{location}, \text{scale}, \text{shape}\}$, which is an exogenous k-states first-order Markov process with the following transition matrix $\boldsymbol{P}^h$:

$$\boldsymbol{P}^h = \begin{bmatrix} p_{1,1}^h & p_{1,2}^h & \cdots & p_{1,k}^h \\ p_{2,1}^h & p_{2,2}^h & \cdots & p_{2,k}^h \\ \vdots & \vdots & \ddots & \vdots \\ p_{k,1}^h & p_{k,2}^h & \cdots & p_{k,k}^h \end{bmatrix}, \tag{6}$$

where $p_{i,j}^h = \Pr(s_t^h = j | s_{t-1}^h = i)$ denotes the transition probability that s_t^h is equal to j given that s_{t-1}^h is equal to i, with $i, j \in \{1, \ldots, k\}$, $p_{i,j}^h \geq 0$ and $\sum_{j=1}^k p_{i,j}^h = 1$.[3] In this way, each parameter can independently change over time according to separate Markov processes. Therefore, the times of changes in one parameter type is stochastically independent of changes in another type, as there is no inherent reason to assume simultaneous changes across all parameters of the distribution. In the online Appendix, I provide evidence that the specification of independent changes is preferred by the data over that of simultaneous changes.[4]

With the introduction of three-independent Markov-switching processes, the overall transition matrix $\boldsymbol{P}$ becomes:

$$\boldsymbol{P} = \boldsymbol{P}^{\text{location}} \otimes \boldsymbol{P}^{\text{scale}} \otimes \boldsymbol{P}^{\text{shape}},$$

with $\otimes$ denotes the Kronecker product.

Although the model is neither linear nor Gaussian, a basic filtering algorithm can be obtained and used as the basis for estimation. This is the objective of the next section.

skew-normal distribution of Azzalini (1985, 1986). See, for example, Genton (2004) for more details.

[3] A natural extension would be to allow transition probabilities to vary over time using exogenous explanatory variables (see Filardo, 1994). Although this approach can be fruitful, it is beyond the scope of the present work.

[4] Furthermore, Lhuissier (2022) demonstrates that allowing for three independent Markov shifts in location, scale and shape parameters offers a more accurate description of real GDP growth compared to synchronized-chains models. Similarly, Sims (2001) and Lhuissier and Zabelina (2015) find that allowing independent transitions for variance and coefficient regimes provides the best fit.

3 State-space Representation, Basic Filtering, and Estimation

In this section, I discuss my model from a state-space perspective, including filtering and estimation.

3.1 State-space Representation

My model is trivially cast in state-space form with regime-switching as

$$\mathbf{y}_t = \mathbf{H}_{s_t}\mathbf{x}_t + \boldsymbol{\varepsilon}_t, \qquad t = 1, \ldots, T, \tag{7}$$

$$\mathbf{x}_t = \mathbf{F}_{s_t}\mathbf{x}_{t-1} + \boldsymbol{\eta}_t, \tag{8}$$

$$\boldsymbol{\varepsilon}_t \sim \text{normal}\left(\boldsymbol{\mu}_{s_t,\varepsilon}, \boldsymbol{\Sigma}_{s_t,\varepsilon}\right), \tag{9}$$

$$\boldsymbol{\eta}_t \sim \text{closed skew-normal}\left(\boldsymbol{\mu}_{s_t,\eta}, \boldsymbol{\Sigma}_{s_t,\eta}, \boldsymbol{\Gamma}_{s_t,\eta}, \boldsymbol{\nu}_{s_t,\eta}, \boldsymbol{\Delta}_{s_t,\eta}\right), \tag{10}$$

$$p_{i,j} = \Pr(s_t = j | s_{t-1} = i), \qquad i, j = 1, \ldots, K, \tag{11}$$

where $\mathbf{y}_t$ is an $N \times 1$ vector of observed variables, $\mathbf{x}_t$ is an $J \times 1$ vector of state variables, $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ are vectors of measurement and transition shocks, containing ς_{it} and ϑ_t. In the online Appendix, I outline precisely the state space form of the proposed DFM.

The model described by equations (7) to (11) becomes the skewed state-space model with Markov-switching. The model includes, as special cases, the Kim (1994)'s Gaussian Markov-switching state-space model if $\boldsymbol{\Gamma}_{s_t,\eta} = \mathbf{0}$, $\boldsymbol{\nu}_{s_t,\eta} = \mathbf{0}$ and $\boldsymbol{\Delta}_{s_t,\eta} = \boldsymbol{I}$, and the constant-parameters skewed state-space model proposed by Naveau, Genton, and Shen (2005), Rezaie and Eidsvik (2014), and Guljanov, Mutschler, and Trede (2025) when the index s_t is dropped from parameters.

3.2 Basic filtering

Suppose the parameters of the model specified in the previous section are known. Let $\boldsymbol{\xi}_{t-1} = \left(\mathbf{y}'_{t-1}, \mathbf{y}'_{t-2}, \ldots, \mathbf{y}'_1\right)'$ denotes the vector of observations available as of time $t-1$. In the standard derivation of the skewed Kalman filter for a fixed-coefficient state-space model (Naveau, Genton, and Shen, 2005; Rezaie and Eidsvik, 2014), the objective is to construct a predictive distribution of the unobserved state vector $\mathbf{x}_t$ based on $\boldsymbol{\xi}_{t-1}$, denoted $\mathbf{x}_{t|t-1}$. In the skewed state-space model with Markov switching, the goal is to characterize the predictive distribution of $\mathbf{x}_t$ not only based on $\boldsymbol{\xi}_{t-1}$, but also conditioned on the random variable s_t taking on the value j and on s_{t-1} taking on the value i:[5]

$$\mathbf{x}_{t|t-1}^{(i,j)} \sim \text{closed skew-normal}\left(\boldsymbol{\mu}_{t|t-1}^{(i,j)}, \boldsymbol{\Sigma}_{t|t-1}^{(i,j)}, \boldsymbol{\Gamma}_{t|t-1}^{(i,j)}, \boldsymbol{\nu}_{t|t-1}^{(i,j)}, \boldsymbol{\Delta}_{t|t-1}^{(i,j)}\right), \tag{12}$$

where

$$\boldsymbol{\mu}_{t|t-1}^{(i,j)} = \boldsymbol{F}_j \boldsymbol{\mu}_{t-1|t-1}^{i} + \boldsymbol{\mu}_{j,\eta}, \qquad \boldsymbol{\Sigma}_{t|t-1}^{(i,j)} = \boldsymbol{F}_j \boldsymbol{\Sigma}_{t-1|t-1}^{i} \boldsymbol{F}_j' + \boldsymbol{\Sigma}_{j,\eta},$$

$$\boldsymbol{\Gamma}_{t|t-1}^{(i,j)} = \begin{pmatrix} \boldsymbol{\Gamma}_{t-1|t-1}^{i} \boldsymbol{\Sigma}_{t-1|t-1}^{i} F_j' \left(\boldsymbol{\Sigma}_{t|t-1}^{(i,j)}\right)^{-1} \\ \boldsymbol{\Gamma}_{j,\eta} \boldsymbol{\Sigma}_{j,\eta} \left(\boldsymbol{\Sigma}_{t|t-1}^{(i,j)}\right)^{-1} \end{pmatrix}, \qquad \boldsymbol{\nu}_{t|t-1}^{(i,j)} = \begin{pmatrix} \boldsymbol{\nu}_{t-1|t-1}^{i} \\ \boldsymbol{\nu}_{j,\eta} \end{pmatrix},$$

$$\boldsymbol{\Delta}_{t|t-1}^{(i,j)} = \begin{pmatrix} \boldsymbol{\Delta}_{t|t-1}^{11} & \boldsymbol{\Delta}_{t|t-1}^{12} \\ \left(\boldsymbol{\Delta}_{t|t-1}^{12}\right)' & \boldsymbol{\Delta}_{t|t-1}^{22} \end{pmatrix},$$

[5] Instead, I could have derived the distribution conditional on ξ_{t-1}, $s_t = j$, $s_{t-1} = i$ and $s_{t-2} = h$ $(h, i, j = 1, 2, \ldots, K)$ to obtain more accurate inferences. In this case, the superscripts (i, j) in equation (12) would be changed to (h, i, j). However, when only s_t shows up in the state-space representation, conditioning on (i, j) is usually enough.

with

$$
\begin{aligned}
\boldsymbol{\Delta}^{11}_{t|t-1} =& \boldsymbol{\Delta}^{i}_{t-1|t-1} + \boldsymbol{\Gamma}^{i}_{t-1|t-1}\boldsymbol{\Sigma}^{i}_{t-1|t-1}\left(\boldsymbol{\Gamma}^{i}_{t-1|t-1}\right)' \\
& - \boldsymbol{\Gamma}^{i}_{t-1|t-1}\boldsymbol{\Sigma}^{i}_{t-1|t-1}\boldsymbol{F}'_{j}\left(\boldsymbol{\Sigma}^{i}_{t|t-1}\right)^{-1}\boldsymbol{F}_{j}\boldsymbol{\Sigma}^{i}_{t-1|t-1}\left(\boldsymbol{\Gamma}^{i}_{t-1|t-1}\right)', \\
\boldsymbol{\Delta}^{22}_{t|t-1} =& \boldsymbol{\Delta}_{j,\eta} + \boldsymbol{\Gamma}_{j,\eta}\boldsymbol{\Sigma}_{j,\eta}\boldsymbol{\Gamma}'_{j,\eta} - \boldsymbol{\Gamma}_{j,\eta}\boldsymbol{\Sigma}_{j,\eta}\left(\boldsymbol{\Sigma}^{(i,j)}_{t|t-1}\right)^{-1}\boldsymbol{\Sigma}_{j,\eta}\boldsymbol{\Gamma}'_{j,\eta}, \\
\boldsymbol{\Delta}^{12}_{t|t-1} =& -\boldsymbol{\Gamma}^{i}_{t-1|t-1}\boldsymbol{\Sigma}^{i}_{t-1|t-1}\boldsymbol{F}'_{j}\left(\boldsymbol{\Sigma}^{i}_{t|t-1}\right)^{-1}\boldsymbol{\Sigma}_{j,\eta}\boldsymbol{\Gamma}'_{j,\eta}.
\end{aligned}
$$

Note that $\boldsymbol{\beta}^{i}_{t-1|t-1}$ is an inference on $\boldsymbol{\beta}_{t-1}$ based on information up to time $t-1$, given $s_{t-1}=i$ and $\boldsymbol{\beta}^{(i,j)}_{t-1|t-1}$ is an inference on $\boldsymbol{\beta}_{t-1}$ based on information up to time $t-1$, given $s_t=j$ and $s_{t-1}=i$, with $\boldsymbol{\beta}\in\{\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Gamma},\boldsymbol{\nu},\boldsymbol{\Delta}\}$. The prediction step calculates a battery of K^2 predictions for each date t, corresponding to every possible value for i and j. Associated with these predictions, the updating steps become as follows:

$$
\mathbf{x}^{(i,j)}_{t|t} \sim \text{closed skew-normal}\left(\boldsymbol{\mu}^{(i,j)}_{t|t}, \boldsymbol{\Sigma}^{(i,j)}_{t|t}, \boldsymbol{\Gamma}^{(i,j)}_{t|t}, \boldsymbol{\nu}^{(i,j)}_{t|t}, \boldsymbol{\Delta}^{(i,j)}_{t|t}\right), \tag{13}
$$

where

$$
\begin{aligned}
\boldsymbol{\mu}^{(i,j)}_{t|t} &= \boldsymbol{\mu}^{(i,j)}_{t|t-1} + \boldsymbol{\Sigma}^{(i,j)}_{t|t-1}\boldsymbol{H}'_{j}(\boldsymbol{H}_{j}\boldsymbol{\Sigma}^{(i,j)}_{t|t-1}\boldsymbol{H}'_{j} + \boldsymbol{\Sigma}_{j,\varepsilon})^{-1}(\mathbf{y}_t - \boldsymbol{H}_{j}\boldsymbol{\mu}^{(i,j)}_{t|t-1} - \boldsymbol{\mu}_{j,\varepsilon}), \\
\boldsymbol{\Sigma}^{(i,j)}_{t|t} &= \boldsymbol{\Sigma}^{(i,j)}_{t|t-1} - \boldsymbol{\Sigma}^{(i,j)}_{t|t-1}\boldsymbol{H}'_{j}(\boldsymbol{H}_{j}\boldsymbol{\Sigma}^{(i,j)}_{t|t-1}\boldsymbol{H}'_{j} + \boldsymbol{\Sigma}_{j,\varepsilon})^{-1}\boldsymbol{H}_{j}\boldsymbol{\Sigma}^{(i,j)}_{t|t-1}, \\
\boldsymbol{\Gamma}^{(i,j)}_{t|t} &= \boldsymbol{\Gamma}^{(i,j)}_{t|t-1}, \\
\boldsymbol{\nu}^{(i,j)}_{t|t} &= \boldsymbol{\nu}^{(i,j)}_{t|t-1} - \boldsymbol{\Gamma}^{(i,j)}_{t|t-1}\boldsymbol{\Sigma}^{(i,j)}_{t|t-1}\boldsymbol{H}'_{j}(\boldsymbol{H}_{j}\boldsymbol{\Sigma}^{(i,j)}_{t|t-1}\boldsymbol{H}'_{j} + \boldsymbol{\Sigma}_{j,\varepsilon})^{-1}(\mathbf{y}_t - \boldsymbol{H}_{j}\boldsymbol{\mu}^{(i,j)}_{t|t-1} - \boldsymbol{\mu}_{j,\varepsilon}), \\
\boldsymbol{\Delta}^{(i,j)}_{t|t} &= \boldsymbol{\Delta}^{(i,j)}_{t|t-1}.
\end{aligned}
$$

Similarly to the standard Kim (1994)'s filter, each iteration of the algorithm above produces an K-fold increase in the number of cases to consider. The key challenge is to collapse the terms to reduce the $(K\times K)$ posteriors $\left(\boldsymbol{\mu}^{(i,j)}_{t|t}, \boldsymbol{\Sigma}^{(i,j)}_{t|t}, \boldsymbol{\Gamma}^{(i,j)}_{t|t}, \boldsymbol{\nu}^{(i,j)}_{t|t}, \quad \text{and} \quad \boldsymbol{\Delta}^{(i,j)}_{t|t}\right)$ into K posteriors $\left(\boldsymbol{\mu}^{j}_{t|t}, \boldsymbol{\Sigma}^{j}_{t|t}, \boldsymbol{\Gamma}^{j}_{t|t}, \boldsymbol{\nu}^{j}_{t|t}, \quad \text{and} \quad \boldsymbol{\Delta}^{j}_{t|t}\right)$ to make the above filter operable. Following Har-

rison and Stevens (1976), Kim (1994) employs weighted averages of terms to collapse them. However, such an approach is not viable for skewed state-space models due to unknown properties of the closed skew-normal distribution. Drawing inspiration from Makov (1979) and Makov and Smith (1980), I therefore employ an alternative approximating method:

$$\boldsymbol{\mu}_{t|t}^{j} = \boldsymbol{\mu}_{t|t}^{(\tilde{s}_{t-1},j)}, \quad \boldsymbol{\Sigma}_{t|t}^{j} = \boldsymbol{\Sigma}_{t|t}^{(\tilde{s}_{t-1},j)}, \quad \boldsymbol{\Gamma}_{t|t}^{j} = \boldsymbol{\Gamma}_{t|t}^{(\tilde{s}_{t-1},j)}, \quad \boldsymbol{\nu}_{t|t}^{j} = \boldsymbol{\nu}_{t|t}^{(\tilde{s}_{t-1},j)}, \quad \text{and} \quad \boldsymbol{\Delta}_{t|t}^{j} = \boldsymbol{\Delta}_{t|t}^{(\tilde{s}_{t-1},j)},$$

where $\tilde{s}_{t-1}$ represents the most likely regime to be in place at time $t-1$, according to the filtered probability estimates. The approximating predictive distribution of $\mathbf{x}_{t|t}$ is now only conditioned on the variable s_t taking on the value j:

$$\mathbf{x}_{t|t}^{j} \sim \text{closed skew-normal} \left(\boldsymbol{\mu}_{t|t}^{j}, \boldsymbol{\Sigma}_{t|t}^{j}, \boldsymbol{\Gamma}_{t|t}^{j}, \boldsymbol{\nu}_{t|t}^{j}, \boldsymbol{\Delta}_{t|t}^{j} \right),$$

which thus allows to evaluate equation (12) for the next period.

Another issue concerns the dimension of the skewness, which increases with each iteration of the filter, as demonstrated by equations (12) and (13). Evaluating cumulative distribution functions of growing dimensions becomes rapidly impractical. Thus, it is necessary to introduce some approximations to ensure the feasibility of the filter. Following Guljanov, Mutschler, and Trede (2025), I reduce the dimensionality of skewness by omitting elements in the cumulative distribution functions that do not dramatically distort the symmetry. For further details, the interested reader is referred to this paper.

3.3 Estimation

The key step in evaluating the overall likelihood function $p(\mathbf{y}_T, \mathbf{y}_{T-1}, \ldots | \boldsymbol{\xi}_0)$ is to obtain the conditional likelihood function at time t:

$$p(\mathbf{y}_t | \boldsymbol{\xi}_{t-1}) = \sum_{j=1}^{K} \sum_{i=1}^{K} p(\mathbf{y}_t | s_{t-1} = i, s_t = j, \boldsymbol{\xi}_{t-1}) p\left(s_{t-1} = i, s_t = j | \boldsymbol{\xi}_{t-1}\right), \tag{14}$$

where $p(\mathbf{y}_t|s_{t-1}=i,s_t=j,\boldsymbol{\xi}_{t-1})$ is the conditional likelihood function, based on information up to time $t-1$, given $s_{t-1}=i$ and $s_t=j$, and is obtained on the prediction error decomposition:

$$\mathbf{y}_t|s_{t-1}=i,s_t=j,\boldsymbol{\xi}_{t-1}\sim\text{closed skew-normal}\left(\tilde{\boldsymbol{y}}_{t|t-1}^{(i,j)},\boldsymbol{\Omega}_{t|t-1}^{(i,j)},\boldsymbol{SK}_{t-1}^{(i,j)},\boldsymbol{\nu}_{t|t-1}^{(i,j)},\boldsymbol{\Xi}_{t|t-1}^{(i,j)}\right),$$

with $\tilde{\boldsymbol{y}}_{t|t-1}^{(i,j)}=\boldsymbol{H}_j\boldsymbol{\mu}_{t|t-1}^{(i,j)}+\boldsymbol{\mu}_{j,\varepsilon}$ is the conditional forecast based on information up to time $t-1$, given $s_{t-1}=i$ and $s_t=j$, $\boldsymbol{\Omega}_{t|t-1}^{(i,j)}=\boldsymbol{H}_j\boldsymbol{\Sigma}_{t|t-1}^{(i,j)}\boldsymbol{H}_j'+\boldsymbol{\Sigma}_{j,\varepsilon}$ is the conditional scale of forecast error; $\boldsymbol{SK}_{t-1}^{(i,j)}=\boldsymbol{\Gamma}_{t|t-1}^{(i,j)}\boldsymbol{GK}_{t-1}^{(i,j)}$ is the skewed Kalman gain based on information up to time $t-1$, given $s_{t-1}=i$ and $s_t=j$ and where $\boldsymbol{GK}_{t-1}^{(i,j)}=\boldsymbol{\Sigma}_{t|t-1}^{(i,j)}\boldsymbol{H}_j'\left(\boldsymbol{H}_j\boldsymbol{\Sigma}_{t|t-1}^{(i,j)}\boldsymbol{H}_j'+\boldsymbol{\Sigma}_{j,\varepsilon}\right)^{-1}$ is the Gaussian Kalman gain up to time $t-1$, given $s_{t-1}=i$ and $s_t=j$; and $\boldsymbol{\Xi}_{t|t-1}^{(i,j)}=\boldsymbol{\Delta}_{t|t-1}^{(i,j)}+\left(\boldsymbol{\Gamma}_{t|t-1}^{(i,j)}-\boldsymbol{SK}_{t-1}^{(i,j)}\boldsymbol{H}_j\right)\boldsymbol{\Sigma}_{t|t-1}^{(i,j)}\left(\boldsymbol{\Gamma}_{t|t-1}^{(i,j)}\right)'$.

The filtering and estimation procedure is completed by making inference on the probability terms. The arguments below follow Hamilton (1989), Kim (1994) and Kim and Nelson (1999) as probability terms remain unchanged in the context of skewed state-space models with regime switching. At the beginning of time t, given the probability of being in regime i, $p(s_{t-1}=i|\boldsymbol{\xi}_{t-1})$, one can calculate:

$$p(s_{t-1}=i,s_t=j|\boldsymbol{\xi}_{t-1})=p_{i,j}p(s_{t-1}=i|\boldsymbol{\xi}_{t-1}),\qquad\text{for}\qquad i,j=1,\ldots,K,$$

where $p_{i,j}$ is the transition probability defined in equation (11). The probability terms can then be updated using information up to time t:

$$p(s_{t-1}=i,s_t=j|\boldsymbol{\xi}_t)=\frac{p(\mathbf{y}_t|s_{t-1}=i,s_t=j,\boldsymbol{\xi}_{t-1})p\left(s_{t-1}=i,s_t=j|\boldsymbol{\xi}_{t-1}\right)}{p(\mathbf{y}_t|\boldsymbol{\xi}_{t-1})}.$$

The remaining probability term can then be calculated as:

$$p(s_t=j|\boldsymbol{\xi}_t)=\sum_{i=1}^{K}p(s_{t-1}=i,s_t=j|\boldsymbol{\xi}_t).$$

Using equation (14), the overall likelihood function (in log units) is then given by:

$$p(\mathbf{y}_T, \mathbf{y}_{T-1}, \ldots | \boldsymbol{\xi}_0) = \sum_{t=1}^{T} \ln \left\{ p(\mathbf{y}_t | \boldsymbol{\xi}_{t-1}) \right\}. \tag{15}$$

In case of Bayesian inference, the overall log-likelihood function in (15) is simply combined with the prior density functions to obtain the posterior density.

To estimate the parameters of the model, a nonlinear optimization procedure such as the `CSMINWEL` program, developed by Christopher A. Sims, can be employed. This procedure aims to maximize the approximate log-likelihood function or the posterior mode (in Bayesian inference) with respect to the unknown parameters of the model. Depending on the dimensionality of the model, complementary methods may be necessary, such as the blockwise optimization method developed by Sims, Waggoner, and Zha (2008). In this approach, a class of richly parameterized multivariate Markov-switching models breaks down parameters into several subblocks. Standard nonlinear optimization routines are then applied to each block, while keeping the other subblocks constant, to maximize the objective function.

If Bayesian inference is employed, after obtaining the posterior mode, a Markov Chain Monte Carlo (MCMC) method is necessary to generate draws from the posterior distribution of the DFM with Markov-switching. A straightforward approach is to use a random walk Metropolis-Hasting (RWMH) algorithm.

In the online Appendix, I report the results of simulation exercises on synthetic data to illustrate my method and assess its efficacy.

4 An Empirical Application

I now present a simple application involving a set of macroeconomic aggregates. I describe in turn the data, the specific variant of the model that I implement, and the adjustments to incorporate the Covid-19 period.

4.1 Data description

Four macroeconomic time series are employed in this paper, which are extracted from the Federal Reserve Economic Data (FRED) covering the period 1959.Q1 to 2024.Q4 ($N = 4$). These variables include: (1) real GDP growth, (2) real personal income less transfer payments, (3) real Manufacturing and trade sales, and (4) total employment (in thousands of persons). Sources and details are reported in the online Appendix.

Data are expressed as one hundred times the first difference of the logarithm of each variable. Data are then standardized by subtracting the sample mean from each variable and dividing by its standard deviation. The resulting series are denoted $y_{i,t}$, with $i = [1, \ldots, 4]$.

4.2 Model Implementation

In the development so far I have allowed for general AR(p) and k^h-regime dynamics. In the empirical model that I now take to the data, I make a simplifying assumption that reduces the number of parameters to be estimated. Specifically, I adopt $\gamma_i(L) = \gamma_i$, for $i = 1, \ldots, 4$, and specify first-order AR specifications for both the common component and the four idiosyncratic components in equations (2) and (4): $\phi(L) = (1 - \phi_1 L)$ and $\psi_i(L) = (1 - \psi_i L)$, where $i = 1, \ldots, 4$. Concerning Markov processes, I assume that $k = 2$, meaning that each process, s_t^h, is governed by a two-states Markov process. This results in a total of $K = 8$ $(= k^3)$ regimes.

Finally, I adopt an identifying assumption since the macroeconomic factor is not identified. In particular, the factor loading γ_1 is normalized to a value of one. This corresponds to the named factor normalization in the DFM literature (e.g., Stock and Watson, 2016).[6] Alternatively, the scale parameter of the first regime, $\sigma_{1,\vartheta}$ could be normalized to a value of one to identify the scale of the index n_t.

[6]The identification issue was initially raised in the context of Gaussian errors. In the online Appendix, I show that this issue is similar in a closed skew-normal environment.

4.3 Covid-19 period

The dataset includes the Covid-19 period, which saw unprecedented levels of volatility in many macroeconomic aggregates. Time series models that do not account for such extreme fluctuations may yield biased or inefficient parameter estimates. There are, at least, three potential approaches to address this issue within my framework.

First, I could introduce a third regime for the variable governing the scale parameter s_t^{scale} to capture the Covid-19 period as an extreme-volatility regime. However, this would increase the total number of regimes to estimate ($K = 12$), resulting in the computation of 144 ($= 12 \times 12$) posteriors $\left(\boldsymbol{\mu}_{t|t}^{(i,j)}, \boldsymbol{\Sigma}_{t|t}^{(i,j)}, \boldsymbol{\Gamma}_{t|t}^{(i,j)}, \boldsymbol{\nu}_{t|t}^{(i,j)}, \quad \text{and} \quad \boldsymbol{\Delta}_{t|t}^{(i,j)}\right)$ in the prediction step defined by equation (12). This approach may become computationally prohibitive due to the large number of evaluations required at each iteration of the filter.

Second, Carriero, Clark, Marcellino, and Mertens (2022) suggest augmenting a standard stochastic volatility model with i.i.d. outliers to accommodate more extreme observations. However, integrating this method into my framework would necessitate a complete revision of the basic filtering procedure outlined in this paper, which exceeds the current scope.

Third, Lenza and Primiceri (2022) propose a method to explicitly model changes in shock volatility by adding scaling factors at the time of volatility change. This approach offers computational simplicity since it leverages precise knowledge of the timing of increased variance in macroeconomic innovations due to the Covid-19. Consequently, I adopt this method in my analysis. Specifically, I modify the distributions of disturbance terms ς_{it} and ϑ_t in equations (3) and (5), respectively, as follows:

$$\varsigma_{it} \sim \text{normal}\left(0, c_t^2 \sigma_{i,\varsigma}^2\right), \quad \text{and} \quad \vartheta_t \sim \text{closed skew-normal}\left(\mu_{s_t^{\text{location}},\vartheta}, c_t^2 \sigma^2_{s_t^{\text{scale}},\vartheta}, \alpha_{s_t^{\text{shape}},\vartheta} 0, 1\right),$$

where c_t is equal to 1 before the time period in which the epidemic begins, which I denote by t^* =2020.Q2. I then assume that $c_{t^*} = c_0$, $c_{t^*+1} = c_1$, $c_{t^*+2} = c_2$, and $c_{t^*+j} = 1 + (c_2 - 1)\rho^{j-2}$, with $[c_0, c_1, c_2, \rho]$ is a vector of unknown coefficients to be estimated. By doing so, I simply

re-scale the scale parameter during the second quarter of 2020 by an unknown parameter c_0, and do the same for the third and fourth quarter of 2020 with two other parameters c_1 and c_2. Thereafter, I assume that the scaling factor decays at a rate ρ. Finally, note also that this re-scaling is common to all shocks of the model.

5 Empirical Results

This section presents the results of the model estimation using Bayesian methods. I begin by describing the prior distributions chosen for the model parameters, including the rationale behind their specification. This is followed by a presentation of the empirical results, with a focus on the posterior distributions of the parameters, estimated using a standard RWMH algorithm.

5.1 Prior

The priors are defined on the left-hand side of Table 1. A few of them deserve further discussion. Regarding the parameters governing the measurement equation, γ_i for $i = \{2, \ldots, 4\}$, I choose a normal prior with a mean of 1.00 and a standard deviation of 1.00. The prior for the autoregressive parameters ϕ_1 and ψ_i also follow the same family of distribution with a mean of 0.00 and a standard deviation of 1.00.

Regarding the prior for the scale parameters $\sigma_{i,\varsigma}$ and $\sigma_{s_t^{\text{scale}},\vartheta}$, I opt for an inverted-gamma distribution with a mean of 0.50 and a standard deviation of 1.00. This choice ensures a wide prior distribution covering a broad parameter space.

The prior for the shape parameter $\alpha_{s_t^{\text{shape}},\vartheta}$ has a normal density with a mean of 0.00 and a standard deviation of 3.00. The prior for the location parameter $\mu_{s_t^{\text{location}},\vartheta}$ also follows a normal distribution with a mean of 0.00 and a standard deviation of 2.00, It may be worth noting that I impose the same prior across all regimes, ensuring that any differences in estimated parameters between regimes are primarily driven by the data (i.e., the likelihood)

rather than the priors.

The prior on the transition matrices, $p_{i,i}^h$ for $i = \{1, 2\}$ and $h \in \{\text{location}, \text{scale}, \text{shape}\}$, follows a beta distribution, with a mean of 0.85 and a standard deviation of 0.15, implying a prior duration of six quarters.

Finally, regarding the parameters related to the Covid-19 period, I follow Lenza and Primiceri (2022) and choose a Pareto distribution for the prior on parameters c_0, c_1, and c_2, with scale and shape parameters both equal to one, so that I can cover a wide range of values. The prior for decay parameter ρ has a beta distribution with a mean of 0.50 and a standard deviation of 0.20.

5.2 Posterior estimates

Estimates presented here are based on samples of 1,000 retained draws, obtained by sampling a total of 11,000 draws, discarding the first 1,000, and retaining every 10th draw of the post-burn sample.

On the right-hand side of Table 1, I report the posterior mode, median, and the 90 percent probability interval for each parameter of the estimated model. A key takeaway is the substantial heterogeneity in parameter estimates across regimes. For example, the median of the location parameter in the first regime is approximately 0.09, whereas it turns negative (–0.24) in the second regime. Regarding scale, the second regime exhibits a value more than three times higher than that of the first regime. The shape parameter is also regime-dependent: it is positive in the first regime ($\alpha_{1,\vartheta} = 8.30$) and negative in the second ($\alpha_{2,\vartheta} = -5.38$). Notably, the 90 percent probability intervals for regime-specific parameters do not overlap, suggesting that these parameters are both well identified and distinct across regimes.

Regarding the posterior probabilities of the transition matrix $\boldsymbol{P}^{\text{location}}$ that governs time variation in the location parameter, there is a difference in persistence across regimes. The posterior median for $p_{1,1}^{\text{location}}$ is 0.94 and that for $p_{2,2}^{\text{location}}$ is 0.97, indicating that the persistence

Table 1: Prior and Posterior Distributions.

	Prior			Posterior			
Coefficient	Density	Para(1)	Para(2)	Mode	Median	[5;	95]
γ_2	N	1.00	1.00	0.9435	0.9323	0.8350	1.0505
γ_3	N	1.00	1.00	1.1029	1.1041	0.9966	1.2193
γ_4	N	1.00	1.00	0.6122	0.5899	0.5248	0.6451
ϕ_1	N	0.00	1.00	0.2899	0.2926	0.1583	0.4736
ψ_1	N	0.00	1.00	−0.2077	−0.2221	−0.3390	−0.1068
ψ_2	N	0.00	1.00	−0.1124	−0.1219	−0.2234	−0.0041
ψ_3	N	0.00	1.00	0.0395	0.0318	−0.0695	0.1473
ψ_4	N	0.00	1.00	0.8343	0.8243	0.7631	0.8807
$\sigma_{1,\varsigma}$	I-G	0.50	1.00	0.4514	0.4458	0.4055	0.4864
$\sigma_{2,\varsigma}$	I-G	0.50	1.00	0.5966	0.5991	0.5684	0.6374
$\sigma_{3,\varsigma}$	I-G	0.50	1.00	0.2327	0.3142	0.1653	0.7943
$\sigma_{4,\varsigma}$	I-G	0.50	1.00	0.1311	0.1406	0.1196	0.1684
$\sigma_{s^{\text{scale}}=1,\vartheta}$	I-G	0.50	1.00	0.3170	0.3135	0.2772	0.3601
$\sigma_{s^{\text{scale}}=2,\vartheta}$	I-G	0.50	1.00	0.8331	0.8382	0.6926	1.0108
$\alpha_{s^{\text{shape}}=1,\vartheta}$	N	0.00	3.00	8.3887	8.2995	5.4104	11.3418
$\alpha_{s^{\text{shape}}=2,\vartheta}$	N	0.00	3.00	−5.3586	−5.3812	−9.0036	−3.2800
$\mu_{s^{\text{location}}=1,\vartheta}$	N	0.00	2.00	0.1490	0.0933	−0.0582	0.2636
$\mu_{s^{\text{location}}=2,\vartheta}$	N	0.00	2.00	−0.2279	−0.2399	−0.3468	−0.1518
$p_{1,1}^{\text{location}}$	B	0.85	0.15	0.9659	0.9392	0.8453	0.9889
$p_{2,2}^{\text{location}}$	B	0.85	0.15	0.9893	0.9657	0.8975	0.9942
$p_{1,1}^{\text{scale}}$	B	0.85	0.15	0.9843	0.9719	0.9497	0.9913
$p_{2,2}^{\text{scale}}$	B	0.85	0.15	0.9675	0.9480	0.8893	0.9851
$p_{1,1}^{\text{shape}}$	B	0.85	0.15	0.9315	0.9249	0.8734	0.9590
$p_{2,2}^{\text{shape}}$	B	0.85	0.15	0.7897	0.7795	0.6471	0.8755
c_0	P	1.00	1.00	22.2550	25.1975	14.9736	42.3350
c_1	P	1.00	1.00	17.2700	12.1338	4.9464	25.7297
c_2	P	1.00	1.00	2.6529	2.2683	1.3312	4.6138
ρ	B	0.60	0.20	0.6126	0.5928	0.2609	0.7841

Note: N stands for Normal, B for Beta, I-G for Inverse-Gamma, and P for Pareto distributions. Para(1) and Para(2) correspond to the means and standard deviations for the N, B, and I-G distributions, and to the scale and shape parameters for P distributions. The 5 percent and 95 percent demarcate the bounds of the 90 percent probability interval. For identification issue, γ_1 is set to 1.

of both regimes is roughly similar (with an average duration of 25 quarters) than the second regime (with an average duration of 33 quarters). A similar pattern is observed for the transition matrix $\boldsymbol{P}^{\text{scale}}$, where the posterior durations of both regimes are almost identical.

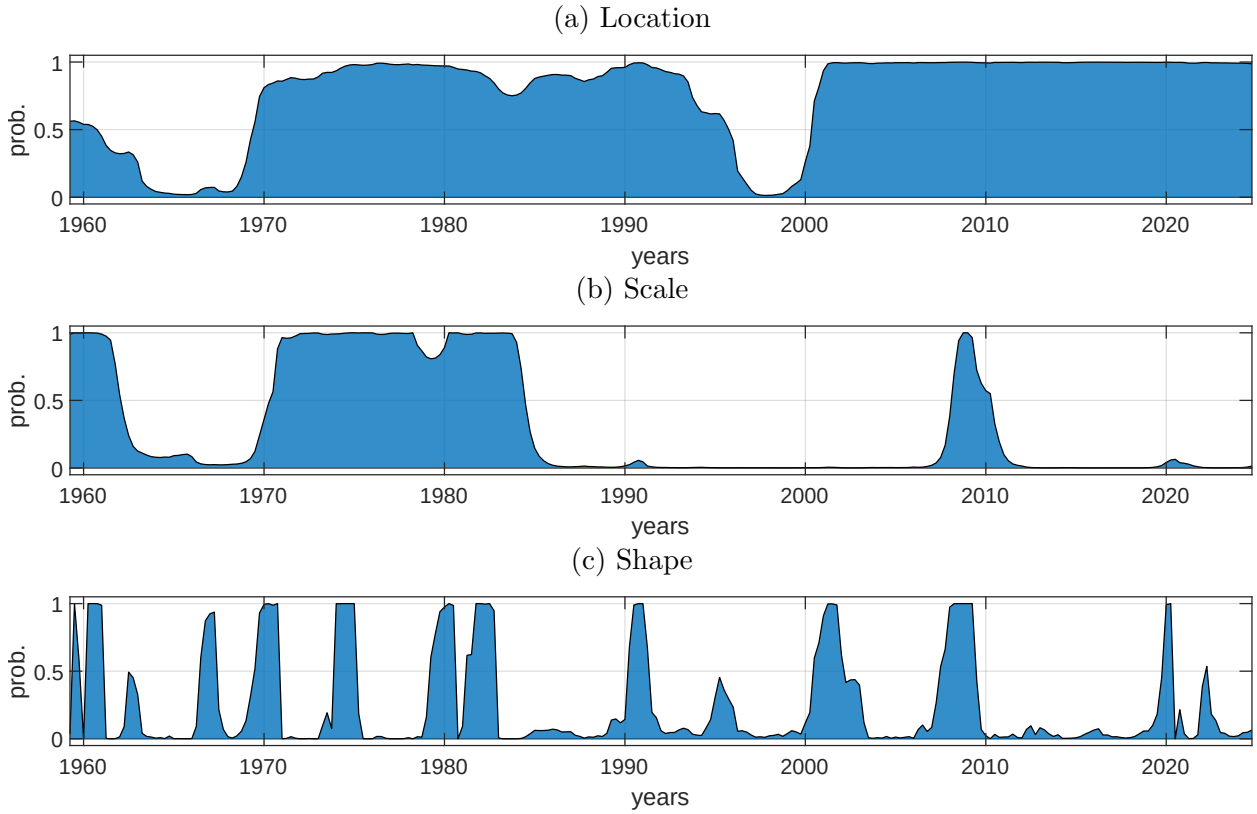
Finally, the transition matrix $\boldsymbol{P}^{\text{shape}}$ reveals an average duration of 5 to 10 quarters for each regime, implying recurrent Markov changes between downside and upside risks in the U.S economy.

Turning to the parameters associated with the Covid-19 episode, the estimates for the scaling coefficients c_0, c_1, and c_2 are well identified, as evidenced by relatively narrow posterior intervals. This indicates that the data are highly informative about these parameters. The estimated value for the rate of decay, ρ, is centered just below 0.60, meaning that Covid-related volatility has been diminishing by 40% each quarter since the fourth quarter of 2020.

Figure 1 reports the regime probabilities, evaluated at the mode, for each Markov process. Probabilities are smoothed in the sense of Kim (1994); i.e., full sample information is used in getting the regime probabilities at each date. One can see from the figure that the U.S economy has been characterized by numerous switches between regimes over time. Looking at the process governing the location parameter (Panel A), the negative-location regime (Regime 2) coincides remarkably well with the 1970s-1980s period, marked by the repeated energy crises that drove up oil costs and weakened U.S growth. This regime has also prevailed since the beginning of 2000s, a period characterized by the slowdown in long-run growth, as typically reported in the literature (e.g., Antolin-Diaz, Drechsel, and Petrella, 2017).

Regarding the process governing the scale parameter (Panel B), the figure indicates that the high-scale regime (Regime 2) predominantly prevailed during the *pre*–Great Moderation period and the 2007-2009 Great Recession. This finding corroborates the works of Sims and Zha (2006), Justiniano and Primiceri (2008), Bianchi (2013), Bianchi and Melosi (2017) and Lhuissier and Tripier (2021), who estimate dynamic time-series models for the U.S economy with volatility changes. After the Great Recession, the high-volatility regime was not anymore evident. It may worth remembering that additional scaling parameters (c_0, c_1, c_2, and ρ) were introduced for the Covid-19 period, and this explains why the Markov-switching process does not capture any regime shifts during that period.

Figure 1: Regime Probabilities



Note: Sample period: 1959.Q2 — 2024.Q4. Evolution of regime probabilities (at the mode) produced from the Markov-switching DFM specification specified by equations (1)-(6). Panels A, B, and C report the probabilities of being in Regime 2 for the three Markov-switching process, s^{location}, s^{scale}, and s^{shape}, respectively. Probabilities are smoothed in the sense of Kim (1994); i.e., full sample information is used in getting the regime probabilities at each date.

Finally, regarding the process governing the shape parameter, Panel C suggests that the U.S economy has experienced numerous switches between negative- and positive-shape regimes over time. Importantly, the negative-shape regime (Regime 2) consistently coincides with recessions, including the 1960–1961 recession, the 1969-1970 recession, the 1973–1975 recession, the recessions of the early 1980s, the early 1990s, and the early 2000s, as well as the Great Recession and the COVID-19 crisis. Although the shape parameter is not the only source of potential asymmetry in the distribution (due to the mixture of distributions), it appears that downside risks are dominant during the episodes of recession. The analysis of the next section will corroborate this finding. Overall, the results clearly show that location, scale, and shape parameters do switch over time, but not in a synchronized manner. This

finding supports the specification of independent Markov-switching processes used in the baseline model.

In order to see whether the results are robust to the choice of the sample period, the model is reestimated excluding the last episode of recession (that is, the Covid period beginning in 2020.Q1). Most of the parameter estimates are similar to the ones reported in Table 1. A modest exception is the estimate of ϕ_1, which is somewhat higher when the post-Covid period is excluded. The plot of the full-sample smoother, using the whole sample, based on the subsample estimates resembles Figure 1. The results are available in the online Appendix.

6 Macroeconomic Skewness Assessment

The previous results demonstrated that all parameters governing the conditional distribution of the macroeconomic factor n_t are time-varying. However, they do not allow for a direct inspection of higher-order moments of the distribution due to the complex nature of the Markov mixture distribution (e.g., Timmermann, 2000; Perez-Quiros and Timmermann, 2001; Bianchi, 2016). For completeness, and in order to provide an accurate characterization of asymmetric macroeconomic risks in the U.S economy, I study in Section 6.1 the evolution of conditional skewness of the macroeconomic factor. In Section 6.2, I compare it with alternative skewness measures drawn from the literature.

6.1 Time variation in macroeconomic skewness

Deriving the conditional moments of a Markov-switching model is not straightforward. Timmermann (2000) characterizes the moments of the basic Markov switching univariate model for cases where the error term follows a t-distribution or a normal distribution. Lhuissier (2022) extends the approach for the case where the error term is a skew-normal distribution. The author shows that the skewness can manifest through two aspects: 1) the degree of asymmetry of the distribution which is governed by the shape parameter; and 2) the mixture

feature of distributions.[7] I complement the approach of Lhuissier (2022) for the case where the error term follows a closed skew-normal distribution, so that I can derive analytically the third moment of the macroeconomic factor, n_t. Technical details are reported in the online Appendix.

Figure 2 presents the time-varying skewness of the unobserved factor, produced from the Markov-switching skewed DFM specified by equations (1) to (6).[8] There is strong evidence of time variation in macroeconomic skewness, which exhibits large swings between positive and negative values throughout the sample. The coefficient of skewness appears clearly to follow a procyclical pattern, which rises during expansions and falls during downturns. Moreover, the skewness tends to decrease prior to the onset of recessions — a feature that corroborates the findings of Delle Monache, De Polis, and Petrella (2024) and that is also observable for the euro area (e.g., Lhuissier, 2022), both of which rely on GDP growth skewness.

In addition to the macroeconomic skewness measure derived from the third central moment reported above, I also compute alternative, quantile-based macroeconomic skewness measures to assess distributional asymmetry without relying on moment assumptions. These non-parametric measures provide robustness to outliers and heavy tails and serve as a useful cross-check. The findings, available in the online Appendix, confirm the procyclicality of macroeconomic skewness.

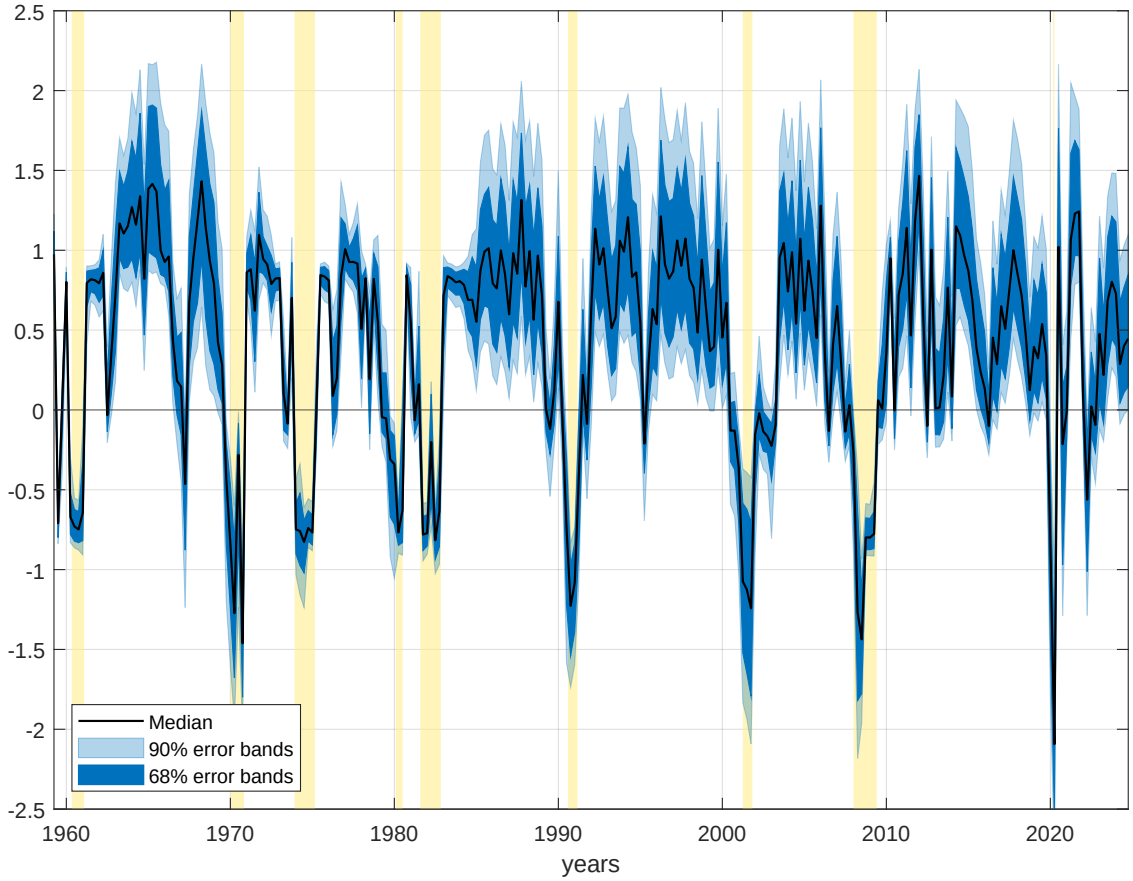
6.2 Comparison with other skewness measures

The results reported in the previous section document strong evidence of time-variation in macroeconomic skewness using a set of four macroeconomic time series. In this subsection I compare it with alternative measures of skewness. Three measures are considered. The first is based on univariate version of the model with GDP growth as measure of economic activity. Since there is no latent factor to extract in this single-variable case, the model

[7] Clark (1973) shows that time-independent mixtures of normal distributions can generate skewness and kurtosis beyond that of a single Gaussian distribution.

[8] The first four moments of the distribution of the macroeconomic factor are presented in the online Appendix.

Figure 2: Macroeconomic Skewness



Note: Sample period: 1959.Q2 — 2024.Q4. Evolution of macroeconomic skewness produced from the Markov-switching DFM specification specified by equations (1)-(6). The median is reported in black solid line and the 68% and 90% error bands in blue areas. The yellow areas denote the NBER-dated recessions.

effectively reduces to an univariate regime-switching skew-normal model. Lhuissier (2022) develops such a model to characterize business cycle variation in the probability distribution of GDP growth in the euro area. Specifically, he focuses on a non-Gaussian model using the skew-normal distribution developed by Azzalini (1985, 1986), where location, scale and shape parameters are allowed to vary over time according to independent two-state Markov-switching processes. Consequently, the Lhuissier (2022)'s univariate Markov-switching model can be seen as a special case of my framework. I apply this methodology to U.S real GDP growth. The data sample starts in the second quarter of 1959 and ends in the fourth quarter of 2019 to avoid the Covid-19 period.

The second measure relies on the methodology developed by Adrian, Boyarchenko, and Giannone (2019). The authors employ a two-step method to estimate flexible parametric predictive distributions of U.S real GDP growth from 1973.Q1 to 2015.Q4. In the first step, distributions are semi-parametrically estimated using quantile regressions that include past economic and financial conditions. In the second step, estimated quantile distributions are smoothed by interpolating between estimated quantiles using flexible skewed t-distributions. This approach converts empirical quantile distributions into estimated conditional distributions of GDP growth. Moments are then easily retrievable. I replicate the measure and extend it through the end of 2019.

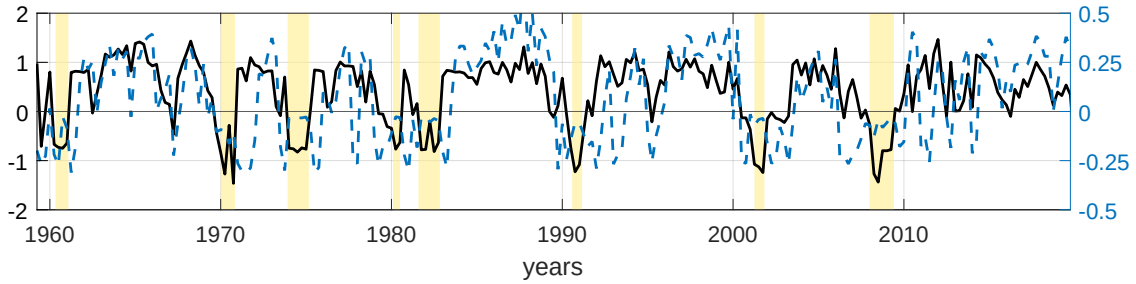
The third measure is drawn from Iseringhausen, Petrella, and Theodoridis (forthcoming). The authors consider a broad range of macroeconomic aggregates to measure macroeconomic skewness. They employ a two-step estimation method, in which series-specific Kelley (1947) skewness measures are estimated by employing autoregressive quantile regressions, prior to extracting their first principal component, which in turn measures the macroeconomic skewness. I obtain the measure until the fourth quarter of 2023 from the authors' website.

Figure 3 plots alternative measures of skewness, along with my estimated macroeconomic skewness measure. Although my measure is positively correlated with the GDP growth skewness, with a correlation coefficient around 0.38, some disparities exist (Panel a). For instance, during the 2001 recession and the Great Recession, my measure exhibits large and persistent declines, accurately capturing heightened downside risks to macroeconomic outcomes. In contrast, the GDP growth skewness declines only marginally during these episodes and remains only slightly negative. An other important episode where the two measures disagree includes the recent years, where my measure of macroeconomic skewness was low and negative but the GDP growth measure of skewness was comparatively high and positive.

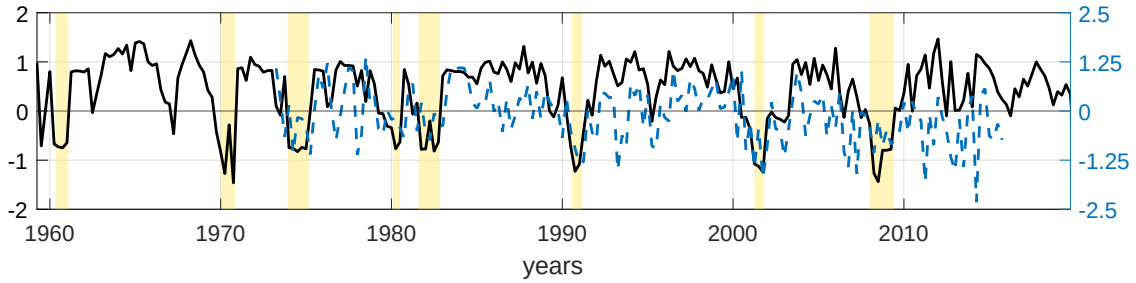
The contrast is also observed with the skewness measure constructed using quantile regressions that include financial conditions (Panel b). It actually suggests frequently and

Figure 3: Macroeconomic Skewness versus Alternative Skewness Measures

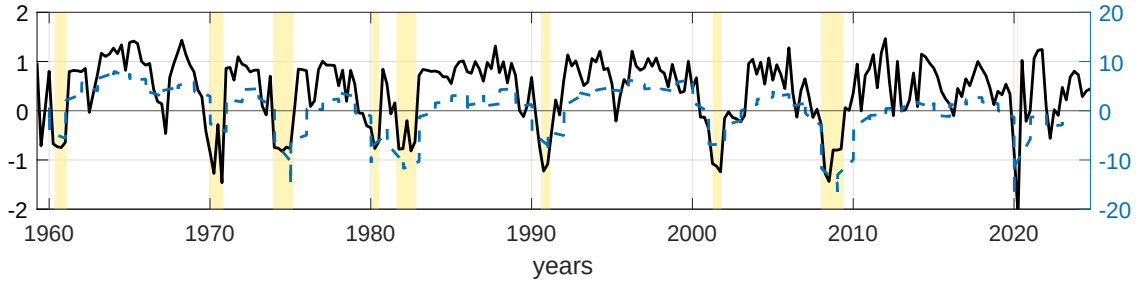
(a) GDP growth skewness (univariate version of the model)



(b) Adrian, Boyarchenko, and Giannone (2019)'s GDP growth skewness



(c) Iseringhausen, Petrella, and Theodoridis (forthcoming)'s macroeconomic skewness



Note: The top-hand panel (Panel a) shows the evolution of a skewness measure based exclusively on GDP growth using the methodology developed by Lhuissier (2022). The middle-hand panel (Panel b) shows the evolution of a skewness measure based on the methodology developed by Adrian, Boyarchenko, and Giannone (2019). The bottom-hand panel (Panel c) shows the evolution of a macroeconomic skewness measure based on the methodology developed by Iseringhausen, Petrella, and Theodoridis (forthcoming). For all panels, the black solid line reports the evolution of the macroeconomic skewness, while the blue dotted line displays that of alternative skewness measures. The yellow areas denote the NBER-dated recessions.

relatively low and negative levels of skewness since the end of the Great Recession in 2009, when the macroeconomic skewness is mostly positive. Another important discrepancy arises in the 1990s, where the skewness of GDP growth which conditions on financial conditions was null. Again, this contrasts with macroeconomic skewness which is at a record high during this period. That being said, both series share some similarities — namely that, they

both fall into negative territory during recessions. This suggests that the heightened asymmetry during these episodes of economic tumult is not necessarily driven by a tightening of financial conditions, and thus corroborates the findings of Plagborg-Møller, Reichlin, Ricco, and Hasenzagl (2020).

A further comparison is drawn with the measure developed by Iseringhausen, Petrella, and Theodoridis (forthcoming), which yields a correlation of approximately 0.50 with my measure of macroeconomic skewness (Panel C). Although the two measures are positively related, my measure is notably more volatile, and episodes of positive skewness occur far more frequently (Panel c). Despite this difference in volatility, both measures point to the same qualitative conclusion: macroeconomic skewness is procyclical, tending to fall into negative territory during recessions. The higher volatility of my measure can be attributed to its construction based on the third central moment, which is highly sensitive to extreme values. In contrast, the measure used by Iseringhausen, Petrella, and Theodoridis (forthcoming) is based on the Kelley skewness, a quantile-based approach that is more robust to outliers and yields smoother dynamics. As discussed above, when I derive the Kelley skewness from my model, the resulting measure is substantially less volatile and closely tracks the dynamics reported in their study.

7 Out-of-Sample Forecasting

In this section, I evaluate the out-of-sample performance of the proposed model in assessing risks to growth. Specifically, I examine whether incorporating a broad set of macroeconomic indicators through my framework yields improvements in predictive accuracy relative to univariate benchmark models.

To assess the information content of the baseline model — referred to as MS-SDFM (Markov-switching Skewed Dynamic Factor Model) — I compare its predictive distributions against two benchmark approaches: 1) A univariate version of the model, which relies

solely on real GDP growth as a summary measure of economic activity (referred to as MS-Univariate); 2) The quantile regression approach of Adrian, Boyarchenko, and Giannone (2019), which models the conditional distribution of GDP growth as a function of lagged macroeconomic and financial indicators (referred to as ABG (2019)).

The out-of-sample procedure is implemented in a pseudo-real-time fashion since it relies on final revised data. The forecasting strategy follows a standard recursive approach. Specifically, I begin with an initial estimation sample spanning 1959:Q2 to 1991.Q4, from which I generate predictive distributions for 1992:Q1 (one quarter ahead) and 1992:Q4 (one year ahead). I then expand the estimation window by one quarter at a time, recursively repeating the estimation and forecasting steps until the end of the sample in 2019.Q4.[9] At each iteration, the estimation follows the methodology detailed in Section 3. Details about the forecasting procedure are available in the online Appendix.

7.1 Downside Risk Assessment

I begin by examining the lower tail of the GDP growth distribution to assess downside risks. As a summary metric, I use the GaR at the 5th percentile, which corresponds to the value below which future GDP growth is expected to fall with 5% probability. This measure is directly obtained from the estimated conditional predictive distributions and provides a clear, interpretable signal of macroeconomic risk.[10]

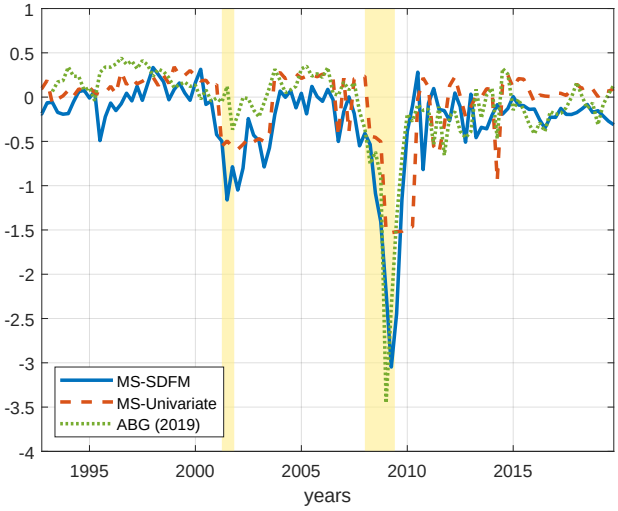
The left panel of Figure 4 plots the 5th percentile GaR estimates across the three models. The baseline model (MS-SDFM) captures key episodes of rising downside risk ahead of U.S. recessions, although only partially. For instance, the MS-SDFM model signals a deterioration in growth prospects ahead of the 2001 recession: the GaR declines to around -0.5% in 2000.Q4 — just before the official onset of the recession — and reaches a trough of approximately -1.0% in early 2001. A similar pattern is observed during the Great Recession: the GaR falls

[9] I do not include the Covid-19 period because the Lenza and Primiceri (2022)'s methodology is not suitable for out-of-sample exercises.

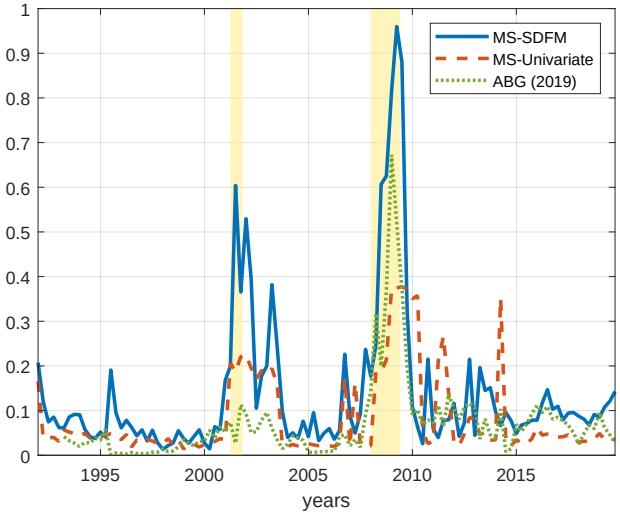
[10] As a complement, I also compute the expected shortfall, which gives the expected value of GDP growth conditional on it being in the worst 5% of outcomes; results are reported in the Appendix.

Figure 4: Downside Risk Assessment.

(a) Growth-at-Risk (5%)



(b) Probability of Recession



Note: Sample period: 1992.Q1 — 2019.Q4. Evolution of the 5% GaR and the probability of recession at the one-quarter ahead horizon for the three models: MS-SDFM, MS-Univariate, and ABG (2019). The yellow areas denote the NBER-dated recessions.

to -0.5% ahead of the 2008 recession and declines further to about -3.0% at the peak of the downturn.

The MS-SDFM's GaR estimates exhibit distinct dynamics compared to the univariate benchmarks. In the case of the 2001 recession, the ABG (2019) model fails to signal rising downside risks — the GaR remains essentially flat throughout the period. This is likely due to the model's dependence on financial conditions, which changed little in the lead-up to that recession, resulting in minimal variation in its conditional distribution of GDP growth. In contrast, the Markov-switching univariate model shows a moderate decline in GaR ahead of the 2001 recession, somewhat echoing the MS-SDFM pattern, though the deterioration is notably less pronounced. During the Great Recession, the ABG (2019) model performs more closely to the MS-SDFM, with both models capturing the sharp decline in GaR as financial conditions deteriorate rapidly. Overall, these results highlight the added value of the MS-SDFM in tracking evolving downside risks — whether or not financial conditions are the main source of economic stress.

Inference about downside risks can also be made by examining the probability of a recession, defined here as the probability of negative GDP growth in the following quarter. These probabilities are shown in the right panel of Figure 4. The MS-SDFM model captures increases in recession risk ahead of economic downturns, although the rise is often modest and does not always signal an imminent recession with high confidence. Nonetheless, the model produces reasonably aligned peaks with NBER-defined recession dates. For example, the probability of recession rises to approximately 50% during the 2001 recession and exceeds 90% during the Great Recession. Importantly, the estimated probabilities drop sharply shortly after each recession ends, reflecting timely updates in the GDP growth conditional distribution.

Compared to the univariate benchmarks, the MS-SDFM provides a more informative and responsive assessment of recession risk. During the 2001 recession, the ABG (2019) model's estimated probability remains flat, reflecting its dependence on financial conditions — which, as discussed earlier, did not deteriorate significantly during that episode. The MS-univariate model does respond somewhat, but its recession probability peaks at just 20%, far below that of the MS-SDFM. During the Great Recession, both univariate models again lag behind: neither generates recession probabilities that match the magnitude or timing of those produced by the MS-SDFM. These results underscore the advantage of incorporating a broad set of macroeconomic signals, as done in the MS-SDFM, for tracking shifts in recession risk.

7.2 Forecast Accuracy

In this section, I perform a more formal evaluation of out-of-sample predictions. I rely on three types of analysis. First, I evaluate the accuracy of the density forecasts. Density forecasts accuracy is evaluated via the predictive log-score. To compute the score with the baseline model (MS-SDFM), I use a kernel-smoothed estimate of the density of the draws from the predictive distribution based on linear diffusion processes as in Botev, Grotowski,

and Kroese (2010). Second, I consider the quantile score, also known as the tick loss function (e.g., Giacomini and Komunjer, 2005), to evaluate the lower tail quantile estimate, at the 5 percent quantile. This metric allows me to provide an assessment of tail risk predictions. Third, I provide an analysis of the calibration of the predictive distribution by looking at the properties of the probability integral transforms (PITs).

Table 2 reports the out-of-sample average log predictive scores and quantile scores for both one-quarter-ahead and one-year-ahead forecasts. To facilitate interpretation, all results for the baseline model (MS-SDFM) are presented relative to those of the univariate benchmark models. Specifically, log score differentials are computed as the difference between the MS-SDFM and the benchmark models, such that positive values indicate superior predictive accuracy of the MS-SDFM. Quantile scores, by contrast, are reported as ratios, with values below 1 indicating better performance of the MS-SDFM relative to the benchmarks. The p-values from the Diebold and Mariano (1995)-West (1996) tests are shown in parentheses, and values significant at the 10% level are highlighted in bold.

Across the full sample, the MS-SDFM shows modest improvements over the MS-Univariate model (Panel A), with positive log score differentials and quantile score ratios below 1 at both horizons, although these differences are not statistically significant. Relative to ABG (2019) (Panel B), the MS-SDFM performs similarly on average, with small and statistically insignificant differences in both metrics.

The model's strength becomes more evident during recessions. The MS-SDFM significantly outperforms both benchmark models in recession periods, especially in terms of log scores. Against the MS-Univariate benchmark, log score gains are large and statistically significant at both horizons, with improvements of 0.91 and 0.76 (p-values < 0.05). The quantile score ratios during recessions also indicate sharper accuracy in capturing tail risks, both for the one-quarter-ahead forecast (ratio = 0.42, p = 0.0999) and the four-quarter-ahead forecast (ratio = 0.59, p = 0.0161). When compared to ABG (2019), the MS-SDFM again displays notable improvements in log scores during recessions, especially at the short-term horizon

Table 2: Out-of-Sample Forecasting Performance Relative to Univariate Benchmarks.

	Panel A. MS-Univariate Model			
	One-quarter ahead		Four-quarter ahead	
	Log-Score	Quantile Score ($\tau = 0.05$)	Log-Score	Quantile Score ($\tau = 0.05$)
Full sample	0.0248	0.7908	0.0286	0.7649
	(0.7184)	(0.1565)	(0.7066)	(0.2057)
Recessions	**0.9055**	**0.4204**	**0.7580**	**0.5945**
	(0.0325)	(0.0999)	(0.0054)	(0.0161)

	Panel B. Adrian, Boyarchenko, and Giannone (2019)'s model			
	One-quarter ahead		Four-quarter ahead	
	Log-Score	Quantile Score ($\tau = 0.05$)	Log-Score	Quantile Score ($\tau = 0.05$)
Full sample	−0.0091	0.8973	−0.0107	1.0833
	(0.9008)	(0.4180)	(0.8852)	(0.4341)
Recessions	**0.7694**	**0.5094**	**0.4733**	0.9128
	(0.0618)	(0.0621)	(0.0809)	(0.5430)

Note: To facilitate interpretation, all results for the baseline model (MS-SDFM) are presented relative to those of the univariate benchmark models (Panel A: MS-Univariate model; Panel B: ABG (2019)). Specifically, log score differentials are computed as the difference between the MS-SDFM and the benchmark models, such that positive values indicate superior predictive accuracy of the MS-SDFM. Quantile scores, by contrast, are reported as ratios, with values below 1 indicating better performance of the MS-SDFM relative to the benchmarks. The p-values from the Diebold and Mariano (1995)-West (1996) tests are shown in parentheses, and values significant at the 10% level are highlighted in bold.

(0.77, p = 0.0618). Quantile score ratios also suggest better downside risk capture, although the four-quarter-ahead improvement is less pronounced and not statistically significant.

Overall, these results underscore the MS-SDFM's superior ability to forecast adverse economic conditions, particularly around turning points, and highlight the value of its broader information set.

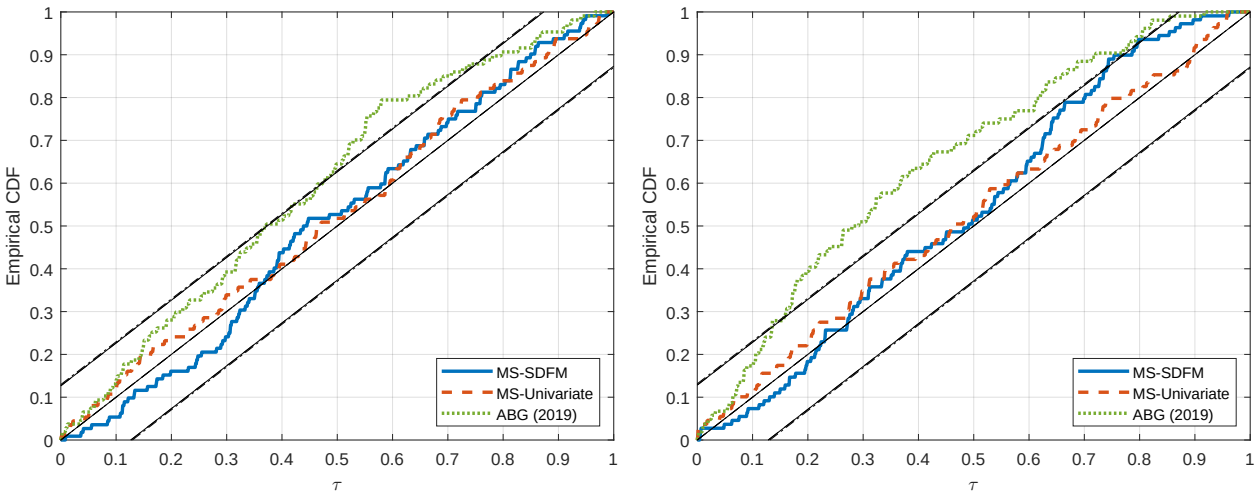
Figure 5 presents the empirical cumulative distribution functions (CDFs) of the PITs for the three models under consideration. Under correct density calibration, the PITs should follow a uniform distribution, implying that the empirical CDF should closely align with the 45-degree line. To formally assess calibration, I construct confidence bands around the 45-degree line using the methodology of Rossi and Sekhposyan (2019), which allows for visual inference about significant deviations from uniformity.

The results indicate that the MS-SDFM exhibits strong calibration. Its empirical PIT

distributions lie entirely within the confidence bands and remain close to the 45-degree line across most of all quantiles and both forecast horizons. The MS-Univariate model also demonstrates good calibration, with PITs tracking the uniform distribution reasonably well. Notably, for the four-quarter-ahead forecasts, it appears slightly better aligned with the 45-degree line at higher quantiles compared to the MS-SDFM. In contrast, the ABG (2019) model shows signs of miscalibration. At the one-quarter-ahead horizon, its PIT CDF deviates from the 45-degree line at certain quantiles, falling outside the confidence bands. These discrepancies become more pronounced at the four-quarter-ahead horizon, where significant deviations are observed across a broader range of quantiles. Overall, these results highlight the MS-SDFM's good calibration performance, particularly in comparison to the ABG (2019) model.

Figure 5: Probability Integral Transforms (PITs) — Real GDP Growth.

(a) One-quarter ahead (b) Four-quarter ahead



Note: Calibration of the predictive distribution by means of the probability integral transforms (PITs). The closer the empirical cumulative distribution of the PITs is to the 45 degrees line, the better the model is calibrated. Critical values are obtained as in Rossi and Sekhposyan (2019) to account for sample uncertainty.

8 Conclusion

In this paper, I introduced a novel framework to assess broad-based, asymmetric macroeconomic risk. By moving beyond GDP growth alone and capturing the joint dynamics of multiple indicators, the proposed model provides a richer view of the balance of risks in the economy. Macroeconomic skewness reveals meaningful cyclical variation and improves the detection of downside risks, particularly around recessions. Forecast evaluations confirm that incorporating a broader information set enhances predictive accuracy, especially in recession periods.

This work can be expanded in several directions. First, a natural extension would be to relax the assumption of exogenous regime switching. This would allow to better understand the causes of shifts in macroeconomic skewness. Second, the provided methodology offers potentially a feasible and reliable way to model and estimate time variation in the skewness of structural shocks within DSGE models. These extensions would open up interesting avenues for further research and could be applied to a variety of economic issues.

References

Adrian, T., N. Boyarchenko, and D. Giannone (2019): "Vulnerable Growth," *American Economic Review*, 109(4), 1263–1289.

Adrian, T., F. Grinberg, N. Liang, S. Malik, and J. Yu (2022): "The Term Structure of Growth-at-Risk," *American Economic Journal: Macroeconomics*, 14(3), 283–323.

Antolin-Diaz, J., T. Drechsel, and I. Petrella (2017): "Tracking the Slowdown in Long-Run GDP Growth," *The Review of Economics and Statistics*, 99(2), 343–356.

——— (2024): "Advances in nowcasting economic activity: The role of heterogeneous dynamics and fat tails," *Journal of Econometrics*, 238(2), 105634.

Azzalini, A. (1985): "A Class of Distributions Which Includes the Normal Ones," *Scandinavian Journal of Statistics*, 12(2), 171–178.

——— (1986): "Further Results on a Class of Distributions Which Includes the Normal Ones," *Statistica*, 46, 199–208.

Barro, R. J. (2009): "Rare Disasters, Asset Prices, and Welfare Costs," *American Economic Review*, 99(1), 243–64.

Barro, R. J., and J. F. Ursúa (2012): "Rare Macroeconomic Disasters," *Annual Review of Economics*, 4(1), 83–109.

Bekaert, G., E. Engstrom, and A. Ermolov (forthcoming): "Uncertainty and the Economy: The Evolving Distributions of Aggregate Supply and Demand Shocks," *American Economic Journal: Macroeconomics*.

Bianchi, F. (2013): "Regimes Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics," *Review of Economic Studies*, 80(2), 463–490.

——— (2016): "Methods for measuring expectations and uncertainty in Markov-switching models," *Journal of Econometrics*, 190(1), 79–99.

Bianchi, F., and L. Melosi (2017): "Escaping the Great Recession," *American Economic Review*, 107(4), 1030–58.

Botev, Z. I., J. F. Grotowski, and D. P. Kroese (2010): "Kernel density estimation via diffusion," *The Annals of Statistics*, 38(5).

Busch, C., D. Domeij, F. Guvenen, and R. Madera (2022): "Skewed Idiosyncratic Income Risk over the Business Cycle: Sources and Insurance," *American Economic Journal: Macroeconomics*, 14(2), 207–42.

Cai, Y., and P. A. Guerrón-Quintana (2023): "Lopsided Interest Rates in International Borrowing Markets," .

Caldara, D., D. Cascaldi-Garcia, P. Cuba-Borda, and F. Loria (2022): "Understanding Growth-at-Risk: A Markov Switching Approach," mimeo.

Caldara, D., H. Mumtaz, and M. Zhong (2024): "Risk in a Data-Rich Model," mimeo.

Caldara, D., C. Scotti, and M. Zhong (2024): "Macroeconomic and Financial Risks: A Tale of Mean and Volatility," mimeo.

Carriero, A., T. E. Clark, and M. Marcellino (2024): "Capturing Macro-Economic Tail Risks with Bayesian Vector Autoregressions," *Journal of Money, Credit and Banking*, 56(5), 1099–1127.

Carriero, A., T. E. Clark, M. Marcellino, and E. Mertens (2022): "Addressing COVID-19 Outliers in BVARs with Stochastic Volatility," *The Review of Economics and Statistics*, pp. 1–38.

Castelnuovo, E., and L. Mori (2025): "Uncertainty, Skewness, and the Business Cycle Through the MIDAS Lens," *Journal of Applied Econometrics*, 40(1), 89–107.

Chauvet, M. (1998): "An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching," *International Economic Review*, 39(4), 969–996.

Ciccarelli, M., and B. Mojon (2010): "Global Inflation," *The Review of Economics and Statistics*, 92(3), 524–535.

Clark, P. K. (1973): "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41(1), 135–155.

Del Negro, M., and C. Otrok (2008): "Dynamic factor models with time-varying parameters: measuring changes in international business cycles," Staff Reports 326, Federal Reserve Bank of New York.

Delle Monache, D., A. De Polis, and I. Petrella (2024): "Modeling and Forecasting Macroeconomic Downside Risk," *Journal of Business & Economic Statistics*, 42(3), 1010–1025.

Diebold, F. X., and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13(3), 253–263.

Filardo, A. J. (1994): "Business-Cycle Phases and Their Transitional Dynamics," *Journal of Business & Economic Statistics*, 12(3), 299–308.

Forni, M., L. Gambetti, and L. Sala (2024): "Downside and Upside Uncertainty Shocks," *Journal of the European Economic Association*, 23(1), 159–189.

Gabaix, X. (2012): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *The Quarterly Journal of Economics*, 127(2), 645–700.

Genton, M. G. (2004): *Skew-elliptical distributions and their applications: a journey beyond normality*. CRC Press.

Giacomini, R., and I. Komunjer (2005): "Evaluation and Combination of Conditional Quantile Forecasts," *Journal of Business & Economic Statistics*, 23, 416–431.

Giglio, S., B. Kelly, and S. Pruitt (2016): "Systemic risk and the macroeconomy: An empirical evaluation," *Journal of Financial Economics*, 119(3), 457–471.

González-Farías, G., J. A. Domínguez-Molina, and A. K. Gupta (2004): *The Closed Skew-Normal Distribution*pp. 25–42. London: Chapman & Hall/CRC, New York.

Gourio, F. (2012): "Disaster Risk and Business Cycles," *American Economic Review*, 102(6), 2734–66.

Guerrón-Quintana, P., A. Khazanov, and M. Zhong (2024): "Nonlinear Dynamic Factor Models," Discussion paper, mimeo.

Guljanov, G., W. Mutschler, and M. Trede (2025): "Pruned skewed Kalman filter and smoother with application to DSGE models," Cqe working papers, Center for Quantitative Economics (CQE), University of Muenster.

Hamilton, J. D. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–384.

Harrison, P. J., and C. F. Stevens (1976): "Bayesian Forecasting," *Journal of the Royal Statistical Society. Series B (Methodological)*, 38(3), 205–247.

Iseringhausen, M., I. Petrella, and K. Theodoridis (forthcoming): "Aggregate Skewness and the Business Cycle," *Review of Economics and Statistics*.

Justiniano, A., and G. E. Primiceri (2008): "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 98(3), 604–641.

Kelley, T. (1947): *Fundamentals of Statistics*, no. n° 2 in Fundamentals of Statistics. Harvard University Press.

Kim, C.-J. (1993): "Unobserved-Component Time Series Models with Markov-Switching Heteroscedasticity: Changes in Regime and the Link between Inflation Rates and Inflation Uncertainty," *Journal of Business & Economic Statistics*, 11(3), 341–349.

——— (1994): "Dynamic Linear Models with Markov-switching," *Journal of Econometrics*, 60, 1–22.

Kim, C.-J., and C. R. Nelson (1999): *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, vol. 1 of *MIT Press Books*. The MIT Press.

Kim, M.-J., and J.-S. Yoo (1995): "New index of coincident indicators: A multivariate Markov switching factor model approach," *Journal of Monetary Economics*, 36(3), 607–630.

Kose, M. A., C. Otrok, and C. H. Whiteman (2003): "International Business Cycles: World, Region, and Country-Specific Factors," *American Economic Review*, 93(4), 1216–1239.

Lenza, M., and G. E. Primiceri (2022): "How to estimate a vector autoregression after March 2020," *Journal of Applied Econometrics*, 37(4), 688–699.

Lhuissier, S. (2022): "Financial Conditions and Macroeconomic Downside Risks in the Euro Area," *European Economic Review*, 143(C).

Lhuissier, S., and F. Tripier (2021): "Regime-Dependent Effects of Uncertainty Shocks: A Structural Interpretation," *Quantitative Economics*, 12(4).

Lhuissier, S., and M. Zabelina (2015): "On the stability of Calvo-style price-setting behavior," *Journal of Economic Dynamics and Control*, 57(C), 77–95.

Loria, F., C. Matthes, and D. Zhang (2024): "Assessing Macroeconomic Tail Risk," *The Economic Journal*, 135(665), 264–284.

Lucas, R. E. (1977): "Understanding business cycles," *Carnegie-Rochester Conference Series on Public Policy*, 5, 7–29.

Makov, U. E. (1979): "Approximations of unsupervised Bayes learning procedures," in *Bayesian Statistics*, ed. by O. L. R. Jacobs. University Press, Valencia.

Makov, U. E., and A. Smith (1980): "Bayesian Detection and Estimation of Jumps in Linear Systems," in *Analysis and Optimization of Stochastic Systems*, ed. by O. L. R. Jacobs. Academic Press.

Mumtaz, H., and P. Surico (2012): "Evolving International Inflation Dynamics: World and Country-Specific Factors," *Journal of the European Economic Association*, 10(4), 716–734.

Naveau, P., M. G. Genton, and X. Shen (2005): "A skewed Kalman filter," *Journal of Multivariate Analysis*, 94(2), 382–400.

Perez-Quiros, G., and A. Timmermann (2001): "Business cycle asymmetries in stock returns: Evidence from higher order moments and conditional densities," *Journal of Econometrics*, 103(1-2), 259–306.

Plagborg-Møller, M., L. Reichlin, G. Ricco, and T. Hasenzagl (2020): "When is growth at risk?," *Brookings Papers on Economic Activity*.

Rezaie, J., and J. Eidsvik (2014): "Kalman filter variants in the closed skew normal setting," *Computational Statistics & Data Analysis*, 75, 1–14.

Rietz, T. A. (1988): "The equity risk premium a solution," *Journal of Monetary Economics*, 22(1), 117–131.

Rossi, B., and T. Sekhposyan (2019): "Alternative tests for correct specification of conditional predictive densities," *Journal of Econometrics*, 208(2), 638–657.

Salgado, S., F. Guvenen, and N. Bloom (2023): "Skewed Business Cycles," NBER Working Papers 26565, National Bureau of Economic Research, Inc.

Sims, C. A. (2001): "Stability and Instability in U.S. Monetary Policy Behavior," *Manuscript, Princeton University*.

Sims, C. A., D. F. Waggoner, and T. Zha (2008): "Methods for Inference in Large Multiple-equation Markov-switching Models," *Journal of Econometrics*, 146, 255–274.

Sims, C. A., and T. Zha (2006): "Were There Regime Switches in U.S. Monetary Policy?," *American Economic Review*, 96(1), 54–81.

Stock, J., and M. Watson (1991): "A Probability Model of the Coincident Economic Indicators," in *The Leading Economic Indicators: New Approaches and Forecasting Records*, ed. by G. Moore, and K. Lahiri, pp. 63–90. Cambridge University Press.

——— (2016): "Chapter 8 - Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics," vol. 2 of *Handbook of Macroeconomics*, pp. 415–525. Elsevier.

Timmermann, A. (2000): "Moments of Markov switching models," *Journal of Econometrics*, 96(1), 75–111.

West, K. D. (1996): "Asymptotic Inference about Predictive Ability," *Econometrica*, 64(5), 1067–1084.

Wolf, E. (2023): "Estimating Growth at Risk with Skewed Stochastic Volatility Models," mimeo.

Assessing Asymmetric Macroeconomic Risk: Online Appendix

Not for Publication

Stéphane Lhuissier[¶¶]

Note to Readers: This online appendix is intended to complement the main paper and is not self-contained. Readers are encouraged to consult the main text for the full context, methodological framework, and empirical results.

[¶¶]Banque de France, 31, Rue Croix des Petits Champs, DGSEI-DEMFI-POMONE 41-1422, 75049 Paris Cedex 01, FRANCE (Email: stephane.lhuissier@hotmail.com; URL: http://www.stephanelhuissier.eu).

Contents

A Simulation Study

In this section, I conduct a number of simulation studies to assess the performance of the proposed algorithm for estimating skewed state-space models with Markov-switching. I assume that the parameters governing the error terms of the transition equation follow a synchronized two-states Markov-switching process, while the remaining parameters remain constant.

From each simulation study, I generate 500 samples of time series of length $T = 100$, $T = 250$ and $T = 500$. For estimation, I employ a Bayesian approach, and compare both the Markov-switching skewed state-space model and the Markov-switching Gaussian state-space model (Kim, 1994). I estimate the parameters governing the distribution $\boldsymbol{\eta}_t$, namely $\boldsymbol{\mu}_{k,\eta}$, $\boldsymbol{\Sigma}_{k,\eta}$, and $\boldsymbol{\Gamma}_{k,\eta}$, as well as transition probabilities $p_{k,k}$ for $k = \{1, 2\}$. All other parameters are fixed at their true values.

The priors are defined in Table A.1. They are very dispersed and cover a large parameter space, so that the likelihood dominates the information obtained in the posterior.

Table A.1: Prior Distribution.

Coefficient	Name of Coefficient	Density	Para(1)	Para(2)
$[\boldsymbol{\mu}_{k,\eta}]_j$	Location	N	0.00	4.00
$[\boldsymbol{\Sigma}_{k,\eta}]_{jj}$	Scale	I-G	0.50	1.00
$[\boldsymbol{\Gamma}_{k,\eta}]_{jj}$	Shape	N	0.00	10.00
$p_{k,k}$	Transition probability	B	0.85	0.15

Note: N stands for Normal, and B for Beta distributions. Para(1) and Para(2) correspond to the means and standard deviations, respectively. The location parameter $[\boldsymbol{\mu}_{k,\eta}]_j$ indicates the k-th regime and the j-th column of $[\boldsymbol{\mu}_{s_t,\eta}]$; The scale parameter $[\boldsymbol{\Sigma}_{k,\eta}]_{jj}$ indicates the k-th regime and the j-th diagonal element of $[\boldsymbol{\Sigma}_{s_t,\eta}]$; The shape parameter $[\boldsymbol{\Gamma}_{k,\eta}]_{jj}$ indicates the k-th regime and the j-th diagonal element of $[\boldsymbol{\Gamma}_{s_t,\eta}]$.

I consider synthetic data generated from three distinct data-generating processes (DGPs): DGP(1) with 2 regimes, 1 latent state, and 1 observable; DGP(2) with 2 regimes, 2 latent states, and 2 observables; and DGP(3) with 2 regimes, 1 latent state, and 4 observables.

These configurations allow for a systematic assessment across varying levels of model complexity. Overall, I demonstrate the proposed algorithm performs an excellent job.

A.1 DGP(1): 2 regimes, 1 latent state, 1 observable

I consider synthetic data generated from Markov-switching skewed state-space model with 2 regimes, 1 latent state, and 1 observable. The true model parameters are as follows:

$$
\begin{aligned}
&\boldsymbol{F} = 0.8, \quad \boldsymbol{H} = 10, \quad \boldsymbol{\mu}_{\varepsilon} = 0, \quad \boldsymbol{\Sigma}_{\varepsilon} = 1, \\
&\boldsymbol{\mu}_{1,\eta} = -2.30, \quad \boldsymbol{\mu}_{2,\eta} = 1.50, \quad \boldsymbol{\Sigma}_{1,\eta} = 2.00, \quad \boldsymbol{\Sigma}_{2,\eta} = 1.00 \\
&\boldsymbol{\Gamma}_{1,\eta} = 3.00, \quad \boldsymbol{\Gamma}_{2,\eta} = -1.50, \quad \boldsymbol{\nu}_{1,\eta} = \boldsymbol{\nu}_{2,\eta} = 0, \quad \boldsymbol{\Delta}_{1,\eta} = \boldsymbol{\Delta}_{2,\eta} = 1, \\
&p_{1,1} = 0.90, \qquad \text{and} \qquad p_{2,2} = 0.95.
\end{aligned}
$$

Table A.2 presents the results of the first simulation study based on DGP(1). For each length, I report the average value and the 5th and 95th percentiles in square brackets. The Markov-switching skewed state-space model delivers accurate estimates of the true parameters, with performance improving as the sample size increases. In particular, the posterior means of the regime-specific means and variances, such as $\boldsymbol{\mu}_{1,\eta}$ and $\boldsymbol{\Sigma}_{1,\eta}$, converge closely to their true values for $T = 500$, and the percentile intervals narrow considerably. The skewness parameters $\boldsymbol{\Gamma}_{1,\eta}$ and $\boldsymbol{\Gamma}_{2,\eta}$ are also estimated with increasing precision as sample size grows. By contrast, the Gaussian model underestimates regime means and variances consistently and fails to capture asymmetries altogether. The estimated transition probabilities are close to their true values in both models, although the skewed model again performs slightly better overall.

Table A.2: Sampling Distributions — Simulation Study DGP(1)

		MS Skewed State-space			MS Gaussian State-space		
Parameter	True Value	T=100	T=250	T=500	T=100	T=250	T=500
$[\boldsymbol{\mu}_{1,\eta}]$	−2.30	−2.03	−2.23	−2.28	−1.21	−1.26	−1.28
		[−2.57;−0.61]	[−2.49;−1.88]	[−2.46;−2.11]	[−1.54;−0.75]	[−1.48;−1.04]	[−1.43;−1.12]
$[\boldsymbol{\mu}_{2,\eta}]$	1.50	1.19	1.39	1.43	0.86	0.86	0.86
		[0.13;1.87]	[0.71;1.74]	[0.88;1.67]	[0.69;1.04]	[0.77;0.97]	[0.77;0.93]
$[\boldsymbol{\Sigma}_{1,\eta}]$	2.00	2.11	1.95	2.00	0.86	0.70	0.64
		[0.59;4.43]	[0.94;3.13]	[1.32;2.77]	[0.40;1.68]	[0.43;1.05]	[0.46;0.88]
$[\boldsymbol{\Sigma}_{2,\eta}]$	1.00	1.07	1.00	0.96	0.55	0.53	0.53
		[0.45;1.85]	[0.50;1.56]	[0.56;1.33]	[0.38;0.78]	[0.42;0.65]	[0.45;0.62]
$[\boldsymbol{\Gamma}_{1,\eta}]$	3.00	3.63	3.46	3.26	–	–	–
		[−1.02;9.14]	[1.31;7.97]	[1.91;5.33]			
$[\boldsymbol{\Gamma}_{2,\eta}]$	−1.50	−0.98	−1.33	−1.38	–	–	–
		[−3.86;2.23]	[−2.46;0.31]	[−2.04;−0.02]			
$p_{1,1}$	0.90	0.92	0.90	0.90	0.91	0.88	0.87
		[0.82;0.99]	[0.82;0.95]	[0.85;0.94]	[0.79;0.98]	[0.79;0.94]	[0.80;0.93]
$p_{2,2}$	0.95	0.95	0.95	0.95	0.95	0.94	0.93
		[0.89;0.99]	[0.92;0.98]	[0.93;0.97]	[0.88;0.99]	[0.89;0.97]	[0.90;0.96]

Note: Sampling distributions based on Markov-switching skewed state-space models and Markov-switching Gaussian state-space models. Each cell contains the average value and the 90 percent percentiles in brackets.

A.2 DGP(2): 2 regimes, 2 latent states, 2 observables

I consider synthetic data generated from a Markov-switching skewed state-space model with 2 regimes, 2 latent states and 2 observables. The true model parameters are as follows:

$$\boldsymbol{F} = 0.9 \times \begin{bmatrix} \cos(-\varrho) & -\sin(-\varrho) \\ \sin(-\varrho) & \cos(-\varrho) \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\mu}_{\varepsilon} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\Sigma}_{\varepsilon} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\boldsymbol{\mu}_{1,\eta} = \begin{bmatrix} 0.3 \\ -0.1 \end{bmatrix}, \quad \boldsymbol{\mu}_{2,\eta} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \boldsymbol{\Sigma}_{1,\eta} = \begin{bmatrix} 0.6 & 0 \\ 0 & 0.3 \end{bmatrix}, \quad \boldsymbol{\Sigma}_{2,\eta} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.2 \end{bmatrix},$$

$$\boldsymbol{\Gamma}_{1,\eta} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{\Gamma}_{2,\eta} = \begin{bmatrix} 5 & 0 \\ 0 & -3 \end{bmatrix}, \quad \boldsymbol{\nu}_{1,\eta} = \boldsymbol{\nu}_{2,\eta} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\Delta}_{1,\eta} = \boldsymbol{\Delta}_{2,\eta} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$p_{1,1} = 0.90, \qquad \text{and} \qquad p_{2,2} = 0.95,$$

with angle $\varrho = \pi/4$.

Table A.3 reports the results for the second simulation study. Regardless of the sample size, the estimation results for the Markov-switching skewed state-space model using the proposed approximation algorithm exhibit little bias in parameter estimates. The main

difference observed with different sample sizes is that the skewness parameters under Regime 2, $\mathbf{\Gamma}_{2,\eta}$ tend to come closer to their true values when T is larger. Furthermore, the uncertainty around estimates appear to shrink as the sample size increases. Overall, the approximating procedure employed in this paper appears to perform remarkably well.

Table A.3: Sampling Distributions — Simulation Study DGP (2)

		MS Skewed State-space			MS Gaussian State-space		
Parameter	True Value	T=100	T=250	T=500	T=100	T=250	T=500
$[\boldsymbol{\mu}_{1,\eta}]_1$	0.30	0.31	0.31	0.30	0.30	0.31	0.30
		[0.05;0.56]	[0.16;0.46]	[0.20;0.40]	[0.04;0.54]	[0.15;0.46]	[0.19;0.40]
$[\boldsymbol{\mu}_{1,\eta}]_2$	−0.10	−0.11	−0.10	−0.10	−0.10	−0.10	−0.10
		[−0.28;0.08]	[−0.20;−0.00]	[−0.17;−0.03]	[−0.28;0.08]	[−0.21;0.00]	[−0.17;−0.03]
$[\boldsymbol{\mu}_{2,\eta}]_1$	1.00	1.05	1.02	1.01	1.21	1.21	1.21
		[0.91;1.23]	[0.94;1.19]	[0.95;1.08]	[1.16;1.26]	[1.18;1.24]	[1.19;1.24]
$[\boldsymbol{\mu}_{2,\eta}]_2$	2.00	1.91	1.94	1.97	1.71	1.72	1.72
		[1.46;2.15]	[1.67;2.10]	[1.73;2.08]	[1.64;1.79]	[1.67;1.76]	[1.68;1.74]
$[\mathbf{\Sigma}_{1,\eta}]_{11}$	0.60	0.60	0.60	0.60	0.59	0.60	0.60
		[0.30;0.94]	[0.44;0.79]	[0.48;0.71]	[0.33;0.89]	[0.44;0.79]	[0.48;0.71]
$[\mathbf{\Sigma}_{1,\eta}]_{22}$	0.30	0.29	0.29	0.30	0.30	0.30	0.30
		[0.15;0.44]	[0.22;0.38]	[0.24;0.36]	[0.16;0.44]	[0.22;0.38]	[0.25;0.36]
$[\mathbf{\Sigma}_{2,\eta}]_{11}$	0.10	0.09	0.10	0.10	0.05	0.05	0.05
		[0.04;0.15]	[0.05;0.14]	[0.07;0.12]	[0.04;0.07]	[0.04;0.06]	[0.05;0.06]
$[\mathbf{\Sigma}_{2,\eta}]_{22}$	0.20	0.20	0.19	0.20	0.12	0.12	0.12
		[0.10;0.34]	[0.11;0.29]	[0.12;0.27]	[0.08;0.16]	[0.09;0.14]	[0.10;0.13]
$[\mathbf{\Gamma}_{1,\eta}]_{11}$	0.00	−0.00	−0.00	−0.00	–	–	–
		[−0.00;−0.00]	[−0.00;−0.00]	[−0.00;0.00]			
$[\mathbf{\Gamma}_{1,\eta}]_{22}$	0.00	−0.00	−0.00	−0.00	–	–	–
		[−0.00;−0.00]	[−0.00;−0.00]	[−0.00;0.00]			
$[\mathbf{\Gamma}_{2,\eta}]_{11}$	5.00	4.39	4.74	4.91	–	–	–
		[0.04;10.04]	[0.31;7.58]	[3.07;6.63]			
$[\mathbf{\Gamma}_{2,\eta}]_{22}$	−3.00	−2.47	−2.58	−2.81	–	–	–
		[−7.06;3.14]	[−4.76;0.44]	[−4.21;−0.03]			
$p_{1,1}$	0.90	0.90	0.90	0.90	0.90	0.90	0.90
		[0.80;0.97]	[0.84;0.95]	[0.86;0.93]	[0.80;0.97]	[0.84;0.95]	[0.86;0.94]
$p_{2,2}$	0.95	0.95	0.95	0.95	0.95	0.95	0.95
		[0.90;0.99]	[0.92;0.98]	[0.93;0.97]	[0.89;0.98]	[0.92;0.97]	[0.93;0.97]

Note: Sampling distributions based on Markov-switching skewed state-space models and Markov-switching Gaussian state-space models. Each cell contains the average value and the 90 percent percentiles in brackets.

Unlike the Markov-switching skewed state-space model, the estimation results for the Markov-switching Gaussian state-space model show considerable bias in some parameter estimates. For example, the estimated value for the location parameter under Regime 2, $[\boldsymbol{\mu}_{2,\eta}]_1$, considerably overestimates the true value. This bias also appears for the estimated value of the scale parameter under Regime 2, $[\mathbf{\Sigma}_{2,\eta}]_{11}$, which underestimates the true value. This bias is expected because the Gaussian assumption fails to account for skewness, resulting in biased estimates for both the location and the scale parameters.

A.3 DGP(3): 2 regimes, 1 latent state, 4 observables

I consider synthetic data generated from a Markov-switching skewed state-space model consisting of 2 regimes, 1 latent state, 4 observables. The true model parameters are as follows:

$$\boldsymbol{F} = 0.8, \quad \boldsymbol{H} = \begin{bmatrix} 10 \\ 2 \\ 5 \\ -3 \end{bmatrix}, \quad \boldsymbol{\mu}_{\varepsilon} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\Sigma}_{\varepsilon} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\boldsymbol{\mu}_{1,\eta} = -2.30, \quad \boldsymbol{\mu}_{2,\eta} = 1.50, \quad \boldsymbol{\Sigma}_{1,\eta} = 2.00, \quad \boldsymbol{\Sigma}_{2,\eta} = 1.00$$

$$\boldsymbol{\Gamma}_{1,\eta} = 0.00, \quad \boldsymbol{\Gamma}_{2,\eta} = -3.50, \quad \boldsymbol{\nu}_{1,\eta} = \boldsymbol{\nu}_{2,\eta} = 0, \quad \boldsymbol{\Delta}_{1,\eta} = \boldsymbol{\Delta}_{2,\eta} = 1,$$

$$p_{1,1} = 0.90, \qquad \text{and} \qquad p_{2,2} = 0.95.$$

Table A.4 presents the estimation results for the third simulation study. Across all sample sizes, the Markov-switching skewed state-space model using the proposed approximation method continues to yield estimates with minimal bias relative to the true parameter values. Notably, the estimation of the skewness parameter in Regime 2, $\boldsymbol{\Gamma}_{2,\eta}$, improves with larger samples, with the average estimate converging closer to its true value of -3.50 and its associated uncertainty narrowing. Compared to the Gaussian specification, the skewed model recovers the regime-dependent means and variances more accurately, especially for $\boldsymbol{\mu}_{2,\eta}$ and $\boldsymbol{\Sigma}_{2,\eta}$, where the Gaussian model exhibits clear bias. As before, the dispersion of the sampling distributions decreases as T increases, indicating greater precision with more data. These findings further support the effectiveness and reliability of the proposed approximation algorithm under a more complex data-generating process.

Table A.4: Sampling Distributions — Simulation Study DGP(3)

		MS Skewed State-space			MS Gaussian State-space		
Parameter	True Value	T=100	T=250	T=500	T=100	T=250	T=500
$[\boldsymbol{\mu}_{1,\eta}]$	−2.30	−2.34	−2.30	−2.30	−2.28	−2.29	−2.28
		[−2.98;−1.77]	[−2.59;−1.98]	[−2.52;−2.11]	[−2.73;−1.62]	[−2.59;−1.98]	[−2.48;−2.06]
$[\boldsymbol{\mu}_{2,\eta}]$	1.50	1.45	1.49	1.49	0.75	0.76	0.75
		[1.20;1.65]	[1.35;1.61]	[1.39;1.59]	[0.60;0.91]	[0.64;0.86]	[0.68;0.82]
$[\boldsymbol{\Sigma}_{1,\eta}]$	2.00	2.32	2.14	2.07	2.26	2.14	2.05
		[1.02;4.54]	[1.43;3.09]	[1.59;2.65]	[1.06;4.36]	[1.45;3.05]	[1.62;2.55]
$[\boldsymbol{\Sigma}_{2,\eta}]$	1.00	0.97	0.99	0.99	0.40	0.39	0.39
		[0.51;1.49]	[0.67;1.33]	[0.79;1.19]	[0.26;0.57]	[0.31;0.48]	[0.32;0.46]
$[\boldsymbol{\Gamma}_{1,\eta}]$	0.00	−0.00	−0.00	−0.00	–	–	–
		[−0.00;0.00]	[−0.00;0.00]	[−0.00;0.00]			
$[\boldsymbol{\Gamma}_{2,\eta}]$	−3.50	−4.14	−3.64	−3.62	–	–	–
		[−9.74;−1.55]	[−5.42;−2.27]	[−4.81;−2.61]			
$p_{1,1}$	0.90	0.91	0.90	0.90	0.90	0.89	0.89
		[0.81;0.98]	[0.84;0.95]	[0.85;0.94]	[0.77;0.97]	[0.81;0.95]	[0.84;0.93]
$p_{2,2}$	0.95	0.95	0.95	0.95	0.95	0.94	0.95
		[0.90;0.99]	[0.92;0.98]	[0.93;0.97]	[0.89;0.99]	[0.91;0.97]	[0.92;0.97]

Note: Sampling distributions based on Markov-switching skewed state-space models and Markov-switching Gaussian state-space models. Each cell contains the average value and the 90 percent percentiles in brackets.

B Data Description

B.1 Main Data Sources

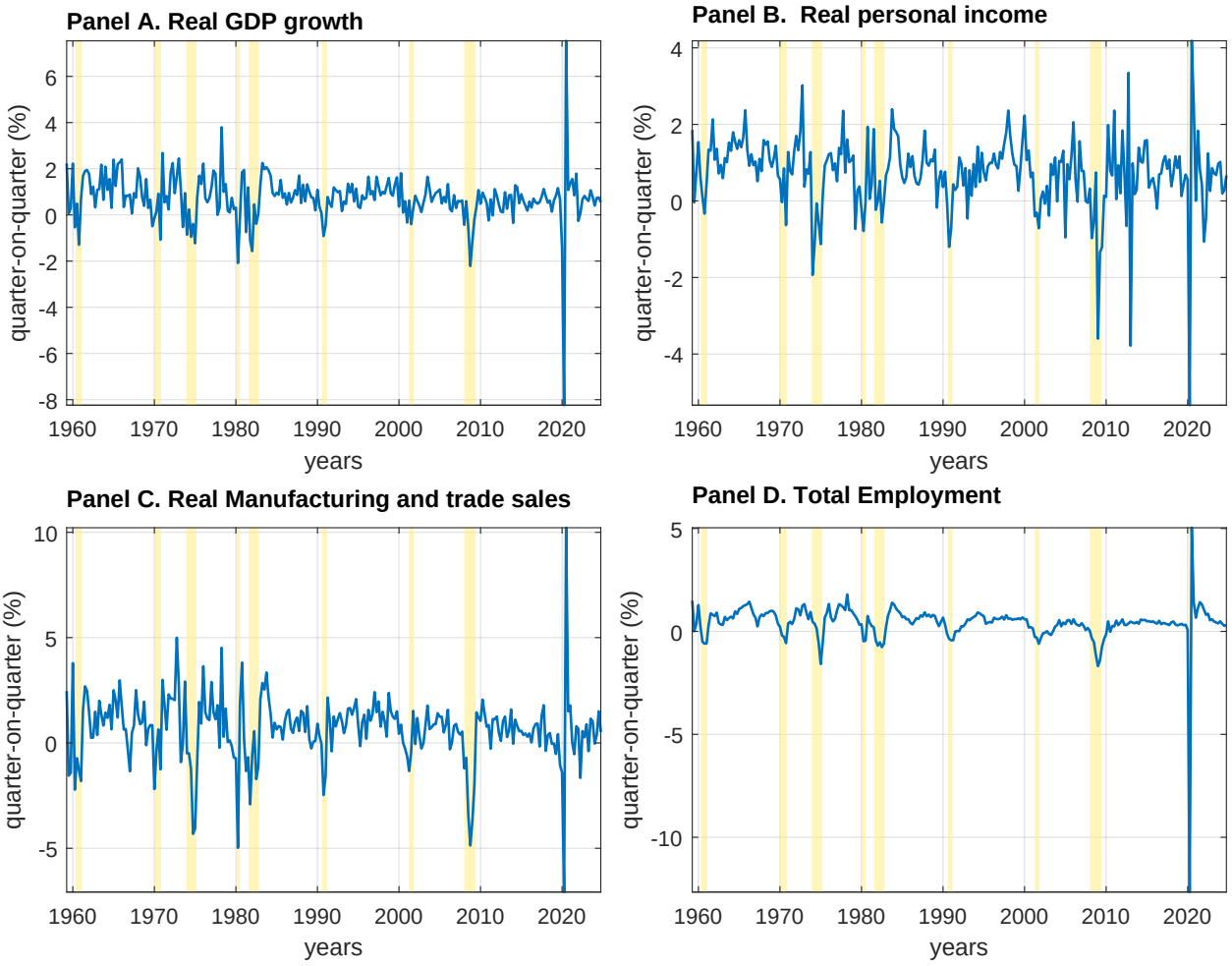
All data are organized quarterly from the first quarter of 1959 to the fourth quarter of 2024. Data comes from the FRED, Federal Reserve Economic Data:

- Real GDP growth: GDPC1
- Real personal income excluding current transfer receipts: W875RX1
- Real Manufacturing and Trade Industries Sales: CMRMTSPL
- All Employees, Total Nonfarm: PAYEMS

Data are expressed as one hundred times the first difference of the logarithm of each variable. Data are then standardized by subtracting the sample mean from each variable and dividing by its standard deviation. Figure B.1 displays the time series data prior to standardization.

B.2 Figure

Figure B.1: Data set for the Dynamic Factor Model with Markov-switching



Note: Sample period: 1959.Q2 — 2024.Q4. The yellow areas denote the NBER-dated recessions.

C The State-Space Form of the Dynamic Factor Model

In this section, I outline the state space form of the dynamic factor model presented in the main manuscript. Consider the following general skewed state-space model with Markov-switching:

$$\mathbf{x}_t = \mathbf{F}_{s_t}\mathbf{x}_{t-1} + \boldsymbol{\eta}_t,$$
$$\mathbf{y}_t = \mathbf{H}_{s_t}\mathbf{x}_t + \boldsymbol{\varepsilon}_t,$$

with $\boldsymbol{\eta}_t$ and $\boldsymbol{\varepsilon}_t$ are assumed to follow closed skew-normal and normal distributions, respectively:

$$\boldsymbol{\eta}_t \sim \text{closed skew-normal}\left(\boldsymbol{\mu}_{s_t,\eta}, \boldsymbol{\Sigma}_{s_t,\eta}, \boldsymbol{\Gamma}_{s_t,\eta}, \boldsymbol{\nu}_{s_t,\eta}, \boldsymbol{\Delta}_{s_t,\eta}\right),$$
$$\boldsymbol{\varepsilon}_t \sim \text{normal}\left(\boldsymbol{\mu}_{s_t,\varepsilon}, \boldsymbol{\Sigma}_{s_t,\varepsilon}\right).$$

Here, $\mathbf{y}_t$ is a vector of contemporaneous endogenous variables, while $\mathbf{x}_t$ is a vector of unobserved state variables. For $1 \leq i, j \leq K$, the discrete and unobserved variable s_t is an exogenous first-order Markov process with the transition matrix $\boldsymbol{P}$:

$$\boldsymbol{P} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,K} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ p_{K,1} & p_{K,2} & \cdots & p_{K,K} \end{bmatrix},$$

where K is the total number of regimes; and $p_{i,j} = \Pr(s_t = j | s_{t-1} = i)$ denotes the transition probability that s_t is equal to j given that s_{t-1} is equal to i, with $p_{i,j} \geq 0$ and $\sum_{j=1}^{K} p_{i,j} = 1$.

The skewed dynamic factor model with Markov-switching, presented in the main manuscript,

can be summarized as follows:

$$
\begin{aligned}
y_{i,t} &= \gamma_i n_t + z_{it}, \qquad i = 1, 2, 3, 4, \\
z_{it} &= \psi_i z_{it-1} + \varsigma_{it}, \qquad \varsigma_{it} \sim \text{normal}\left(0, \sigma^2_{i,\varsigma}\right), \\
n_t &= \phi_1 n_{t-1} + \vartheta_t, \qquad \vartheta_t \sim \text{closed skew-normal}(\mu_{s_t^{\text{location}},\vartheta}, \sigma^2_{s_t^{\text{scale}},\vartheta}, \alpha_{s_t^{\text{shape}},\vartheta}, 0, 1),
\end{aligned}
$$

where the variable $s_t = \left\{s_t^{\text{location}}, s_t^{\text{scale}}, s_t^{\text{shape}}\right\}$ or in a more concise manner $s_t = \left\{s_t^{\text{loc}}, s_t^{\text{sc}}, s_t^{\text{sh}}\right\}$ contains the three unobserved variables that follow independent first-order two-states Markov processes. It implies that the total number of regimes is $K = 8$.

A state-space representation of the model is given by:

$$
\begin{aligned}
\mathbf{x}_t &= \mathbf{F}\mathbf{x}_{t-1} + \boldsymbol{\eta}_t, \\
\mathbf{y}_t &= \mathbf{H}\mathbf{x}_t,
\end{aligned}
$$

where the state, measurement and error variables are as follows:

$$
\mathbf{y}_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \end{bmatrix}, \qquad \mathbf{x}_t = \begin{bmatrix} z_{1t} \\ z_{2t} \\ z_{3t} \\ z_{4t} \\ n_t \end{bmatrix}, \qquad \boldsymbol{\eta}_t = \begin{bmatrix} \varsigma_{1t} \\ \varsigma_{2t} \\ \varsigma_{3t} \\ \varsigma_{4t} \\ \vartheta_t \end{bmatrix},
$$

and transition and measurement equations are not dependent upon the state variable s_t:

$$
\mathbf{F} = \begin{bmatrix} \psi_1 & 0 & 0 & 0 & 0 \\ 0 & \psi_2 & 0 & 0 & 0 \\ 0 & 0 & \psi_3 & 0 & 0 \\ 0 & 0 & 0 & \psi_4 & 0 \\ 0 & 0 & 0 & 0 & \phi_1 \end{bmatrix}, \qquad \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & \gamma_1 \\ 0 & 1 & 0 & 0 & \gamma_2 \\ 0 & 0 & 1 & 0 & \gamma_3 \\ 0 & 0 & 0 & 1 & \gamma_4 \end{bmatrix},
$$

.

The state errors $\boldsymbol{\eta}_t$ follows the Markov-switching closed skew-normal distribution with the following parameters:

$$\boldsymbol{\mu}_{s_t^{\text{lo}}=1,\eta} = \left[0,0,0,0,\mu_{s_t^{\text{lo}}=1,\vartheta}\right]', \qquad \boldsymbol{\mu}_{s_t^{\text{lo}}=2,\eta} = \left[0,0,0,0,\mu_{s_t^{\text{lo}}=2,\vartheta}\right]',$$

$$\boldsymbol{\Sigma}_{s_t^{\text{sc}}=1,\eta} = \text{diag}\left(\left[\sigma_{1,\varsigma},\sigma_{2,\varsigma},\sigma_{3,\varsigma},\sigma_{4,\varsigma},\sigma_{s_t^{\text{sc}}=1,\vartheta}\right]\right), \qquad \boldsymbol{\Sigma}_{s_t^{\text{sc}}=2,\eta} = \text{diag}\left(\left[\sigma_{1,\varsigma},\sigma_{2,\varsigma},\sigma_{3,\varsigma},\sigma_{4,\varsigma},\sigma_{s_t^{\text{sc}}=2,\vartheta}\right]\right),$$

$$\boldsymbol{\Gamma}_{s_t^{\text{sh}}=1,\eta} = \text{diag}\left(\left[0,0,0,0,\alpha_{s_t^{\text{sh}}=1,\vartheta}\right]\right), \qquad \boldsymbol{\Gamma}_{s_t^{\text{sh}}=2,\eta} = \text{diag}\left(\left[0,0,0,0,\alpha_{s_t^{\text{sh}}=2,\vartheta}\right]\right),$$

$$\boldsymbol{\nu}_{s_t,\eta} = \boldsymbol{\nu}_{\eta} = \mathbf{0}_{5\times 1}, \qquad \boldsymbol{\Delta}_{s_t,\eta} = \boldsymbol{\Delta}_{\eta} = \boldsymbol{I}_{5\times 1},$$

where diag is the matrix operation which transforms a vector into a diagonal matrix. The measurement errors $\boldsymbol{\varepsilon}_t$ are dropped from the model.

During the Covid-19 period, the model is adjusted and includes four additional parameters: the three scale parameters c_0, c_1 and c_2, and the decay rate ρ. These modifications to the model are highlighted in red:

$$\boldsymbol{\Sigma}_{s_t^{\text{sc}}=1,\eta} = c_t\text{diag}\left(\left[\sigma_{1,\varsigma},\sigma_{2,\varsigma},\sigma_{3,\varsigma},\sigma_{4,\varsigma},\sigma_{s_t^{\text{sc}}=1,\vartheta}\right]\right), \qquad \boldsymbol{\Sigma}_{s_t^{\text{sc}}=2,\eta} = c_t\text{diag}\left(\left[\sigma_{1,\varsigma},\sigma_{2,\varsigma},\sigma_{3,\varsigma},\sigma_{4,\varsigma},\sigma_{s_t^{\text{sc}}=2,\vartheta}\right]\right),$$

where c_t is equal to 1 before the time period in which the epidemic begins, which I denote by t^* =2020.Q2. I then assume that $c_{t^*} = c_0$, $c_{t^*+1} = c_1$, $c_{t^*+2} = c_2$, and $c_{t^*+j} = 1 + (c_2 - 1)\rho^{j-2}$, with $[c_0, c_1, c_2, \rho]$ the vector of unknown parameters.

The vector of parameters to be estimated by Bayesian inference is as follows:

$$\theta = [\psi_1,\psi_2,\psi_3,\psi_4,\phi_1,\gamma_1,\gamma_2,\gamma_3,\gamma_4,\sigma_{1,\varsigma},\sigma_{2,\varsigma},\sigma_{3,\varsigma},\sigma_{4,\varsigma},\mu_{1,\vartheta},\mu_{2,\vartheta},\sigma_{2,\vartheta},\alpha_{1,\vartheta},\alpha_{2,\vartheta},c_0,c_1,c_2,\rho].$$

D Identification of the Macroeconomic Factor

In this section, I show that there is an identification issue with the baseline model, and an additional normalization is needed. The demonstration presented here is drawn and adapted from a case study by Guerrón-Quintana, Khazanov, and Zhong (2024). This issue is not specific to the non-Gaussian framework of the model and it is similar to that of a Gaussian environment. To prove this, it is convenient to use a constant-parameters version of the model:[12]

$$\begin{aligned} y_{i,t} &= \gamma_i n_t + z_{it}, \qquad i = 1, 2, 3, 4, \\ z_{it} &= \psi_i z_{it-1} + \sigma_{i,\varsigma}\varsigma_{it}, \qquad \varsigma_{it} \sim \text{normal}\,(0, 1)\,, \\ n_t &= \mu_\vartheta + \phi_1 n_{t-1} + \sigma_\vartheta \vartheta_t, \qquad \vartheta_t \sim \text{closed skew-normal}(0, 1, \sigma_\vartheta \alpha_\vartheta, 0, 1). \end{aligned} \tag{16}$$

Consider the following constant $a \neq 0$ and scale the system in equation (16) as follows:

[12] The closed skew-normal distribution is closed under linear transformations. Let $\boldsymbol{X}$ be a d-dimensional random variable as follows $\boldsymbol{X} \sim \text{closed skew-normal}\,(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x, \boldsymbol{\Gamma}_x, \boldsymbol{\nu}_x, \boldsymbol{\Delta}_x)$ and A be a full-rank $d \times d$ matrix, then

$$\boldsymbol{Y} = AX \sim \text{closed skew-normal}\,(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y, \boldsymbol{\Gamma}_y, \boldsymbol{\nu}_y, \boldsymbol{\Delta}_y)\,,$$

with $\boldsymbol{\mu}_y = A\boldsymbol{\mu}_x$, $\boldsymbol{\Sigma}_y = F\boldsymbol{\Sigma}_x F'$, $\boldsymbol{\Gamma}_y = \boldsymbol{\Gamma}_x F^{-1}$, $\boldsymbol{\nu}_y = \boldsymbol{\nu}_x$, and $\boldsymbol{\Delta}_y = \boldsymbol{\Delta}_x$.

Thus, the law of motion of the macroeconomic factor in equation (16) is equivalent to

$$\begin{aligned} n_t &= \mu_\vartheta + \phi_1 n_{t-1} + \vartheta_t, \\ \vartheta_t &\sim \text{closed skew-normal}(0, \sigma^2_\vartheta, \alpha_\vartheta, 0, 1), \end{aligned}$$

and to

$$\begin{aligned} n_t &= \phi_1 n_{t-1} + \vartheta_t, \\ \vartheta_t &\sim \text{closed skew-normal}(\mu_\vartheta, \sigma^2_\vartheta, \alpha_\vartheta, 0, 1). \end{aligned}$$

$$
\begin{aligned}
y_{i,t} &= \gamma_i \left(\frac{1}{a}\right) a n_t + z_{it}, \\
a z_{it} &= \psi_i a z_{it-1} + a \sigma_{i,\varsigma} \varsigma_{it}, \\
a n_t &= a \mu_\vartheta + \phi_1 a n_{t-1} + a \sigma_\vartheta \vartheta_t.
\end{aligned}
$$

Then, define $\tilde{n}_t = a n_t$, $\tilde{\gamma}_i = \gamma_i \left(\frac{1}{a}\right)$, $\tilde{z}_{it} = a z_{it-1}$, $\tilde{\sigma}_{i,\varsigma} = a \sigma_{i,\varsigma}$, $\tilde{\mu}_\vartheta = a \mu_\vartheta$, and $\tilde{\sigma}_\vartheta = a \sigma_\vartheta$, and rewrite the system as follows:

$$
\begin{aligned}
y_{i,t} &= \tilde{\gamma}_i \tilde{n}_t + z_{it}, \\
\tilde{z}_{it} &= \psi_i \tilde{z}_{it-1} + \tilde{\sigma}_{i,\varsigma} \varsigma_{it}, \\
\tilde{n}_t &= \tilde{\mu}_\vartheta + \phi_1 \tilde{n}_{t-1} + \tilde{\sigma}_\vartheta \vartheta_t,
\end{aligned}
\tag{17}
$$

The "tilde" model in equation (17) and the baseline model in equation (16) are observationally equivalent because they imply the same distribution of y_{it} for $i = 1, 2, 3, 4$. The lack of identification is resolved by imposing a mathematically convenient normalization. In particular, the factor loading γ_1 can be normalized to one. This corresponds to the named factor normalization in the DFM literature (e.g., Stock and Watson, 2016). An alternative normalization is to fix the scale parameter to a value of one to identify the scale of the index n_t.

E Model Comparison

The baseline model assumes that the times of changes for a specific parameter are stochastically independent of the times of changes for another one. Here, I relax this assumption and assume that location, scale, and shape parameters switch at the same time. Say it differently, there is only one Markov-switching process (also called "chain") governing all parameters of the model. I compare the fit of this synchronized-chain model with that of the baseline model, and show that the the independent-chains model is preferred.

I employ the Watanabe-Akaike Information Criterion (WAIC), introduced by Watanabe (2010), for purposes of model comparison. WAIC evaluates the predictive accuracy for a fitted model by computing the log pointwise predictive density corrected from the effective number of parameters to adjust for overfitting. WAIC offers two main advantages. First, it is fully Bayesian in that it is based on the usual posterior simulations of the parameters. Second, it is invariant to parametrization.

WAIC is defined as follows:

$$\text{WAIC} = \log(lpd) - \overline{p}, \tag{18}$$

where $\log(lpd)$ is the log pointwise predictive density, i.e., $\sum_{t=1}^{T} \log\left(\frac{1}{S}\sum_{s=1}^{S} p(y_t|\theta^s)\right)$ with $p(y_t|\theta^s)$ is the predictive density at time t conditional on the set of parameters θ^s, and S is the number of MCMC iteration; and $\overline{p}$ is the estimated effective number of parameters, computed based on the posterior variance of the log predictive density for each data point y_t:

$$\overline{p} = \sum_{t=1}^{T} V_{s=1}^{S}(\log(p(y_t|\theta^s)),$$

with V represent the sample variance, $V_{s=1}^{S} a_s = \frac{1}{S-1}(a_s - \bar{a})^2$. It is trivial to compute the standard error by computing the standard deviation of the T components and multiplying

by $\sqrt{T}$. Define

$$\text{WAIC}_t = \log\left(\frac{1}{S}\sum_{s=1}^{S} p(y_t|\theta^s)\right) - \left(V_{s=1}^{S}\log p(y_t|\theta^s)\right),$$

so that WAIC in (18) is the sum of these t independent terms. Then the standard deviation of WAIC is defined as follows

$$se(\text{WAIC}) = \sqrt{(T * V_{t=1}^{T}\text{WAIC}_t)}.$$

When comparing two fitted models, say models A and B, one can estimate the difference in their expected predictive accuracy by the difference in WAIC. The standard error of this difference can be computed as follows:

$$se(\text{WAIC}^A - \text{WAIC}^B) = \sqrt{T * V_{t=1}^{T}(\text{WAIC}_t^A - \text{WAIC}_t^B)}.$$

Table E.1 reports the value and the standard error of WAIC for synchronized- and independent-chains models, as well as their difference. Clearly the specification of independent-chains outperforms that of synchronized-chains as the estimated difference in their expected predictive accuracy is well above zero. This is more clearly appreciated when taking into account the uncertainty (in terms of standard errors) of their difference.

Table E.1: Information criteria

	WAIC	standard errors
Independent-chains model	−868.5482	43.1326
Synchronized-chains model	−875.6439	41.9749
Relative difference	7.0957	5.1583

Note: A model with a higher WAIC is preferred in a model comparison.

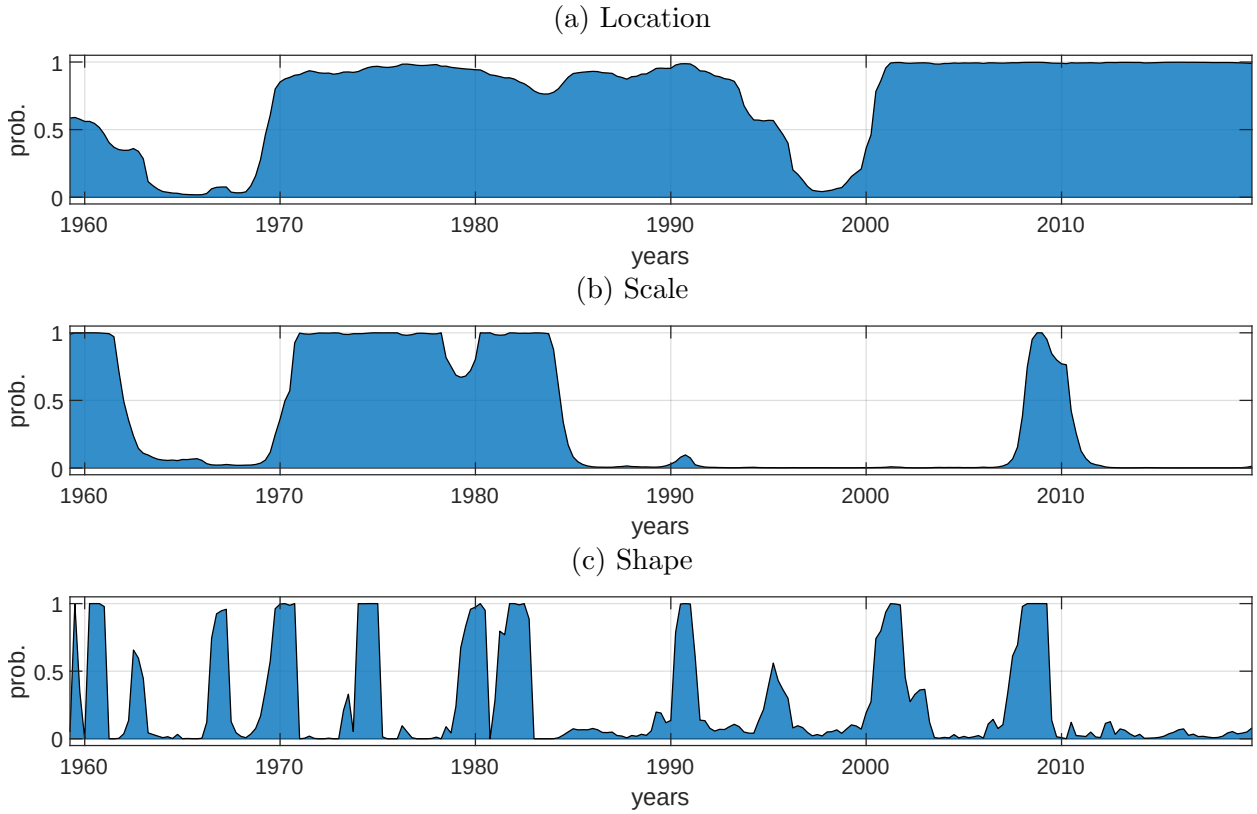
F Results when excluding the post-Covid-19 period

Table F.1: Prior and Posterior Distributions.

	Prior			Posterior			
Coefficient	Density	Para(1)	Para(2)	Mode	Median	[5;	95]
γ_2	N	1.00	1.00	0.8226	0.8179	0.7252	0.9228
γ_3	N	1.00	1.00	0.9666	0.9628	0.8485	1.0811
γ_4	N	1.00	1.00	0.9638	0.9377	0.8506	1.0467
ϕ_1	N	0.00	1.00	0.4665	0.4800	0.3638	0.6409
ψ_1	N	0.00	1.00	−0.2304	−0.2328	−0.3390	−0.1263
ψ_2	N	0.00	1.00	−0.1328	−0.1367	−0.2407	−0.0362
ψ_3	N	0.00	1.00	0.0151	0.0385	−0.0929	0.1772
ψ_4	N	0.00	1.00	0.8411	0.8278	0.7677	0.8888
$\sigma_{1,\varsigma}$	I-G	0.50	1.00	0.5949	0.5941	0.5378	0.6522
$\sigma_{2,\varsigma}$	I-G	0.50	1.00	0.6732	0.6812	0.6425	0.7220
$\sigma_{3,\varsigma}$	I-G	0.50	1.00	0.2323	0.2761	0.1647	0.5919
$\sigma_{4,\varsigma}$	I-G	0.50	1.00	0.2260	0.2369	0.1947	0.2846
$\sigma_{s^{\text{scale}}=1,\vartheta}$	I-G	0.50	1.00	0.3283	0.3242	0.2886	0.3707
$\sigma_{s^{\text{scale}}=2,\vartheta}$	I-G	0.50	1.00	0.9798	0.9827	0.8344	1.1714
$\alpha_{s^{\text{shape}}=1,\vartheta}$	N	0.00	5.00	8.7863	8.1547	1.7042	13.9572
$\alpha_{s^{\text{shape}}=2,\vartheta}$	N	0.00	5.00	−5.2703	−6.7355	−12.8847	−3.1728
$\mu_{s^{\text{location}}=1,\vartheta}$	N	0.00	2.00	0.1680	0.1924	0.0309	0.3186
$\mu_{s^{\text{location}}=2,\vartheta}$	N	0.00	2.00	−0.2463	−0.2487	−0.3886	−0.0243
$p_{1,1}^{\text{location}}$	B	0.85	0.15	0.9640	0.9401	0.8147	0.9892
$p_{2,2}^{\text{location}}$	B	0.85	0.15	0.9879	0.9683	0.8681	0.9961
$p_{1,1}^{\text{scale}}$	B	0.85	0.15	0.9817	0.9737	0.9375	0.9916
$p_{2,2}^{\text{scale}}$	B	0.85	0.15	0.9659	0.9458	0.8859	0.9845
$p_{1,1}^{\text{shape}}$	B	0.85	0.15	0.9269	0.9074	0.8344	0.9542
$p_{2,2}^{\text{shape}}$	B	0.85	0.15	0.7934	0.8038	0.6768	0.9322

Note: N stands for Normal, B for Beta, I-G for Inverse-Gamma, and P for Pareto distributions. Para(1) and Para(2) correspond to the means and standard deviations for the N, B, and I-G distributions, and to the scale and shape parameters for P distributions. The 5 percent and 95 percent demarcate the bounds of the 90 percent probability interval. For identification issue, γ_1 is set to 1.

Figure F.1: Regime Probabilities



Note: Sample period: 1959.Q2 — 2019.Q4. Evolution of regime probabilities (at the mode) produced from the Markov-switching DFM specification specified by equations (1)-(6). Panels A, B, and C report the probabilities of being in Regime 2 for the three Markov-switching process, s^{location}, s^{scale}, and s^{shape}, respectively. Probabilities are smoothed in the sense of Kim (1994); i.e., full sample information is used in getting the regime probabilities at each date.

G The Moments

G.1 Methodology

Timmermann (2000) characterizes the moments of the ergodic distribution of Gaussian Markov-switching models. Perez-Quiros and Timmermann (2001) extend the approach to cover the first four conditional moments. Lhuissier (2022) complements these works for the case where within-regime innovations follow a skew-normal distribution. In this section, I show how to extend the approach of Lhuissier (2022) for the case where innovations have a special form of closed skew-normal distribution.

Consider the following p-order autoregressive skew-normal Markov-switching process:

$$
\begin{aligned}
y_t &= \sum_{j=1}^{p} \varphi_j y_{t-j} + \nu_t, \\
\nu_t &\sim \text{closed skew-normal}(\mu_{s_t}, \sigma_{s_t}^2, \alpha_{s_t}, 0, 1),
\end{aligned}
$$

where φ_j is the j-order autoregressive parameter; and s_t is an exogenous K-states first-order Markov process. The above process is equivalent to

$$
\begin{aligned}
y_t &= \mu_{s_t} + \sum_{j=1}^{p} \varphi_j y_{t-j} + \sigma_{s_t} \varsigma_t, \\
\varsigma_t &\sim \text{closed skew-normal}(0, 1, \hat{\alpha}_{s_t}, 0, 1),
\end{aligned}
$$

where $\hat{\alpha}_{s_t} = \sigma_{s_t} \alpha_{s_t}$ and ς_t is the standardized skew-normal distribution of Azzalini (1985).[13] By expressing ς_t as a skew-normal distribution, I can use the formulas developed by Lhuissier (2022) to compute moments. Recall that $p_{i,t} = \Pr(s_t = i | \boldsymbol{\xi}_{t-1}, \theta)$ is the (filtered) probability of being in regime i, with $i \in \{1, \dots, K\}$, at time t given information at time $t-1$, $\boldsymbol{\xi}_{t-1}$, and define $\tilde{\mu}_{i,t} = \mu_i + \sum_{i=j}^{p} \varphi_j y_{t-j}$. From the law of iterated expectations, the centered moments

[13] The distribution $\varsigma_t \sim \text{closed skew-normal}(0, 1, \hat{\alpha}_{s_t}, 0, 1)$ is equivalent to $\varsigma_t \sim \text{skew-normal}(0, 1, \hat{\alpha}_{s_t})$.

of the process are given by

$$\begin{aligned}
\mathrm{E}[(y_t - \tilde{\mu}_t)^n | \boldsymbol{\xi}_{t-1}, \theta] &= \mathrm{E}\left[\mathrm{E}[(y_t - \tilde{\mu}_t)^n | \boldsymbol{\xi}_{t-1}, \theta, s_t]\right], \\
&= \sum_{i=1}^{K} \mathrm{E}[(\tilde{\mu}_{i,t} + \sigma_{i,t}\varsigma_t - \tilde{\mu}_t)^n | \boldsymbol{\xi}_{t-1}, \theta], \\
&= \sum_{i=1}^{K} p_{i,t} \sum_{j=0}^{n} C_j^n \sigma_{i,t}^j \mathrm{E}(\varsigma_t^j)(\tilde{\mu}_{i,t} - \tilde{\mu}_t)^{n-j},
\end{aligned}$$

where I used Newton's binomial formula. The expression for the case where ς_t follows a normalized skew-normal distribution is based on the moment-generating function (see Section 2.1.4 in Azzalini (2013) for further details). Specifically, it implies that

$$\mathrm{E}[(y_t - \tilde{\mu}_t)^n | \boldsymbol{\xi}_{t-1}, \theta] = \sum_{i=1}^{K} p_{i,t} \sum_{j=0}^{n} C_j^n \sigma_{i,t}^j a_{i,j} (\tilde{\mu}_{i,t} - \tilde{\mu}_t)^{n-j},$$

with $a_{i,j}$ is the j-th raw moment of ς_t conditional on regime i as described below

$$a_{i,j} = \sqrt{\frac{2}{\pi}} \frac{\mathrm{sgn}(\hat{\alpha}_i)}{\hat{\alpha}_i^{j+1}} K_j(\hat{\alpha}_i^{-2}),$$

where

$$K_j(h) = \frac{j-1}{h} K_{j-2}(h) + \frac{\upsilon_{j-1}}{h(1+h)^{j/2}}, \qquad j = 2, 3, \ldots$$

and υ_k is the k-th raw moment of the standard normal distribution, that is,

$$\upsilon_k = \begin{cases} 0 & \text{if } k = 1, 3, 5, \ldots \\ (k-1)!! & \text{if } k = 2, 4, 6, \ldots \end{cases}$$

For $j = 0, 1$, $K_j(h)$ is defined as follows

$$K_0(h) = \left(\frac{\pi}{2h}\right)^{1/2}, \qquad \text{and} \qquad K_1(h) = \frac{1}{h\sqrt{1+h}}.$$

Simplifying the expression $a_{i,j}$ for $j = 0, \ldots, 4$, it follows that

$$
\begin{aligned}
a_{i,0} &= 1, \\
a_{i,1} &= \frac{\hat{\alpha}_i \sqrt{2}}{\sqrt{1 + \hat{\alpha}_i^2} \sqrt{\pi}}, \\
a_{i,2} &= 1, \\
a_{i,3} &= \frac{\left(\hat{\alpha}_i \sqrt{2}\right)\left(3 + 2\hat{\alpha}_i^2\right)}{\sqrt{\pi}(1 + \hat{\alpha}_i^2)^{3/2}}, \\
a_{i,4} &= 3.
\end{aligned}
$$

While the first and second moments, $\mathrm{E}[y_t|\boldsymbol{\xi}_{t-1}, \theta] = \tilde{\mu}_t$ and $\mathrm{E}[(y_t - \tilde{\mu}_t)^2|\boldsymbol{\xi}_{t-1}, \theta]$ are not transformed, I characterize the third ($n = 3$) and fourth ($n = 4$) moments with their corresponding standardized moments, defined respectively as the coefficient of skewness ($\sqrt{b_1}$) and the coefficient of excess kurtosis (b_2) as follows

$$
\sqrt{b_1} \equiv \frac{\mathrm{E}[(y_t - \tilde{\mu}_t)^3]}{\left(\mathrm{E}[(y_t - \tilde{\mu}_t)^2]\right)^{3/2}}, \qquad b_2 \equiv \frac{\mathrm{E}[(y_t - \tilde{\mu}_t)^4] - 3\left(\mathrm{E}[(y_t - \tilde{\mu}_t)^2]\right)^2}{\left(\mathrm{E}[(y_t - \tilde{\mu}_t)^2]\right)^2}.
$$

In the context of the model stated in this paper, one can use the information up to time t to infer the moments.

G.2 Empirical Evidence

Figure G.1 presents the time-varying moments of the distribution of the macroeconomic factor. There is a clear cyclical pattern in the mean of the common factor which declines toward the end of expansions and rise during the end of the recession periods (Panel A). It may be worth mentioning that the decline in the mean during recessions is not only driven by shifts in the location parameter of the model, but also by shifts in the scale and shape parameters, as well as transition probabilities. Thus this explains why the mean declines so

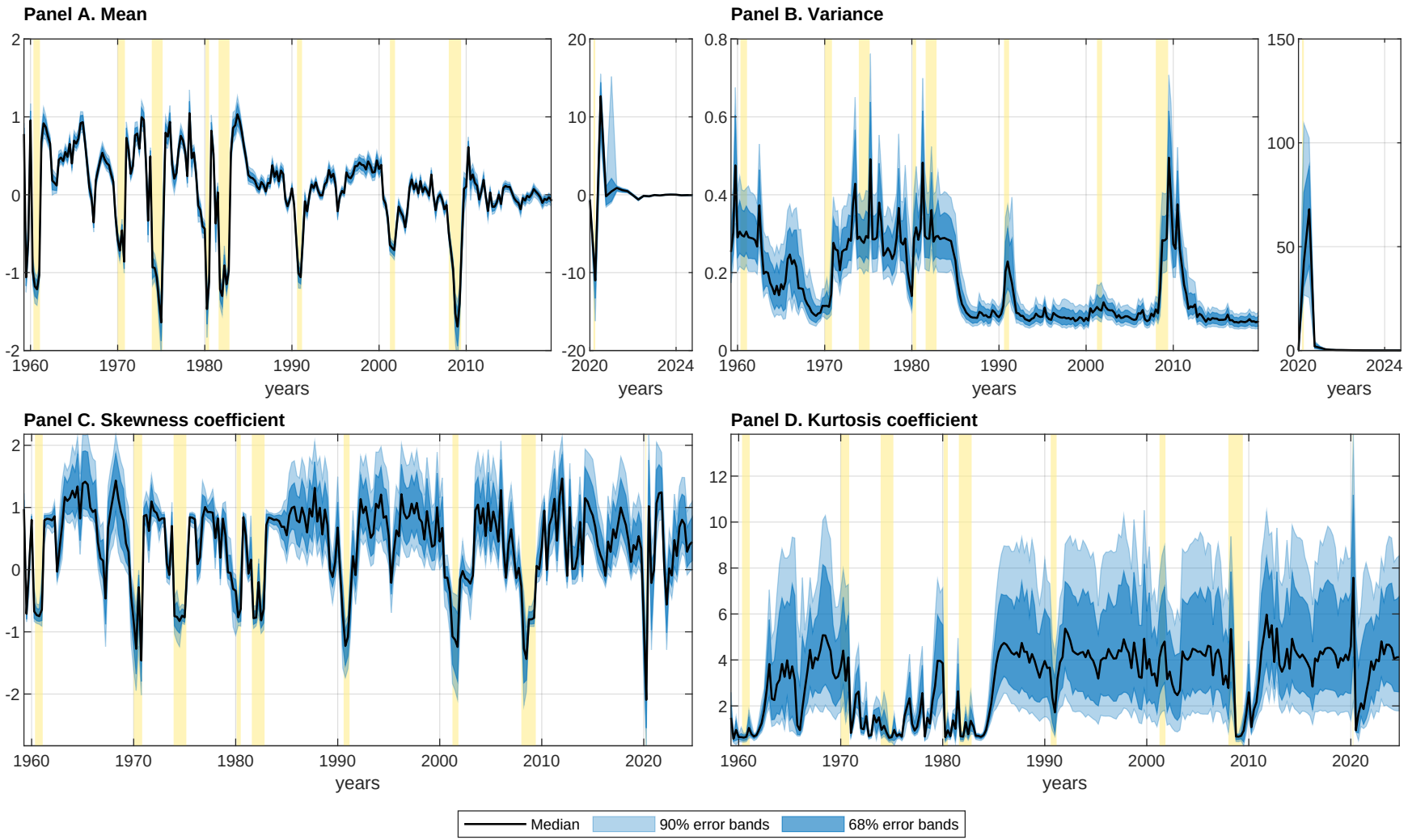
dramatically during the Great Recession, whereas in Figure 1 of the main manuscript, there is no shift in the location parameter during that period.

Turning to the second moment, the common factor generates clear counterfactual patterns, with higher volatility occurring during the pre-Great Moderation, the Great Recession and the Covid-19 crisis. The Covid-related peak is largely driven by estimated scaling factors, which affect the scale parameters of the distributions of the model.

Regarding skewness, the reader should refer to the main manuscript for detailed comments.

Finally, the coefficient of excess kurtosis has a tendancy to rapidly decline during the three recessions, meaning that the distribution of the macroeconomic factor was relatively flat during these episodes. For the rest of the sample, it remains relatively stable, with short-lived oscillations.

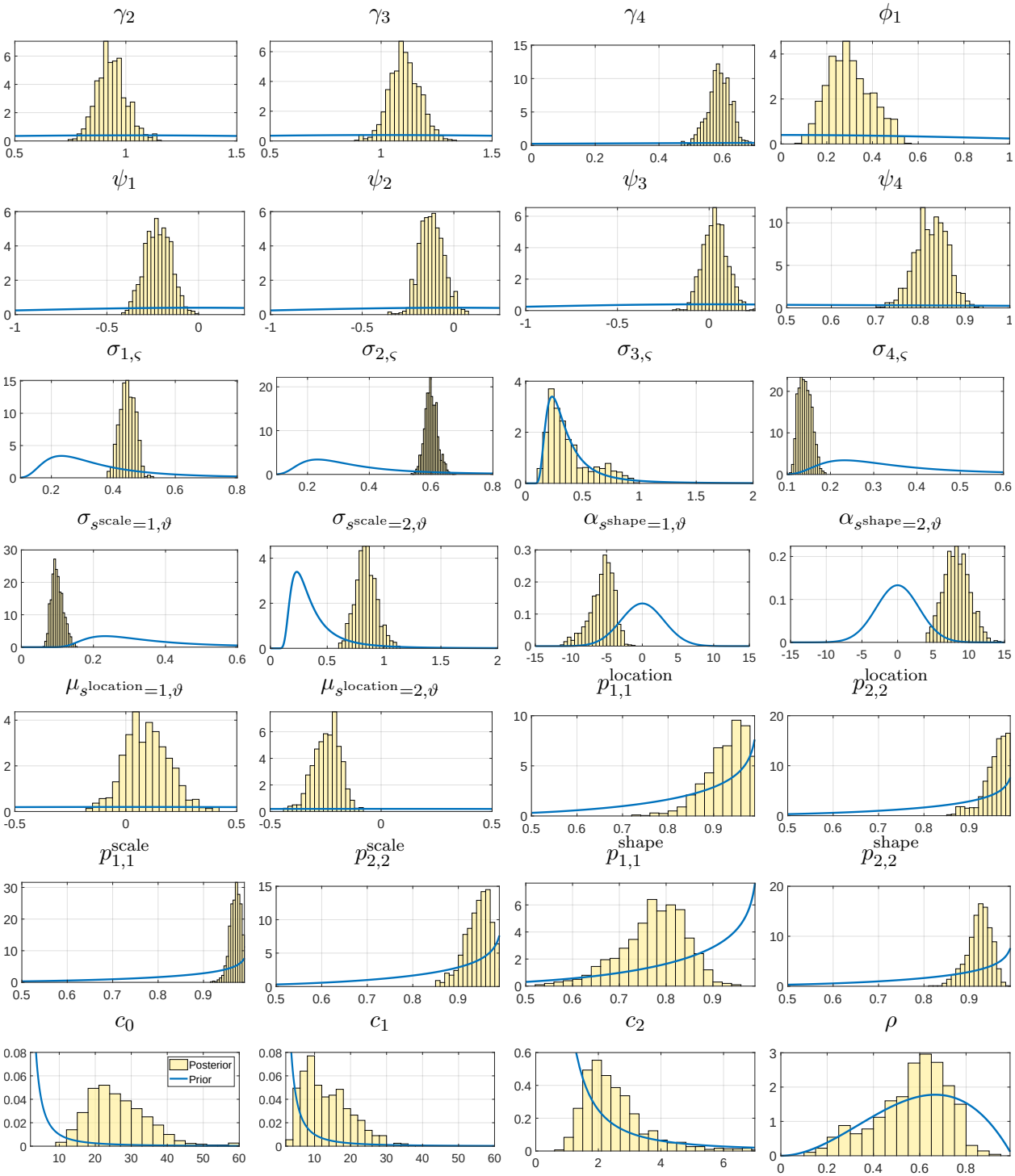
Figure G.1: Moments of the macroeconomic factor, n_t



Note: Sample period: 1959.Q2 — 2024.Q4. Each panel reports the median in black solid line along with the 68% and 90% error bands in blue. The yellow areas denote the NBER-dated recessions.

H Prior and Posterior Distribution of Parameters

Figure H.1: Prior and Posterior Distribution of Parameters



Note: In each panel, the histogram (yellow) represents the posterior distribution, while the blue line depicts the prior distribution. The posterior distribution is obtained using the whole sample.

I Alternative Macroeconomic Skewness Measures based on Quantiles

In addition to the macroeconomic skewness measure derived from the third central moment reported in the main analysis, I also compute alternative, quantile-based macroeconomic skewness measures to assess distributional asymmetry without relying on moment assumptions. These non-parametric measures provide robustness to outliers and heavy tails and serve as a useful cross-check of the results.

Specifically, I implement two versions of Kelley (1947)'s skewness measure, defined as follows:

$$\text{Skewness}_{K(p)} = \frac{Q_{1-p} + Q_p - 2Q_{0.5}}{Q_{1-p} - Q_p},$$

where Q_p denotes the p-th quantile of the distribution. I compute this measure for two symmetric percentile pairs:

- **Kelley(90/10) Skewness**, using the 90th and 10th percentiles:

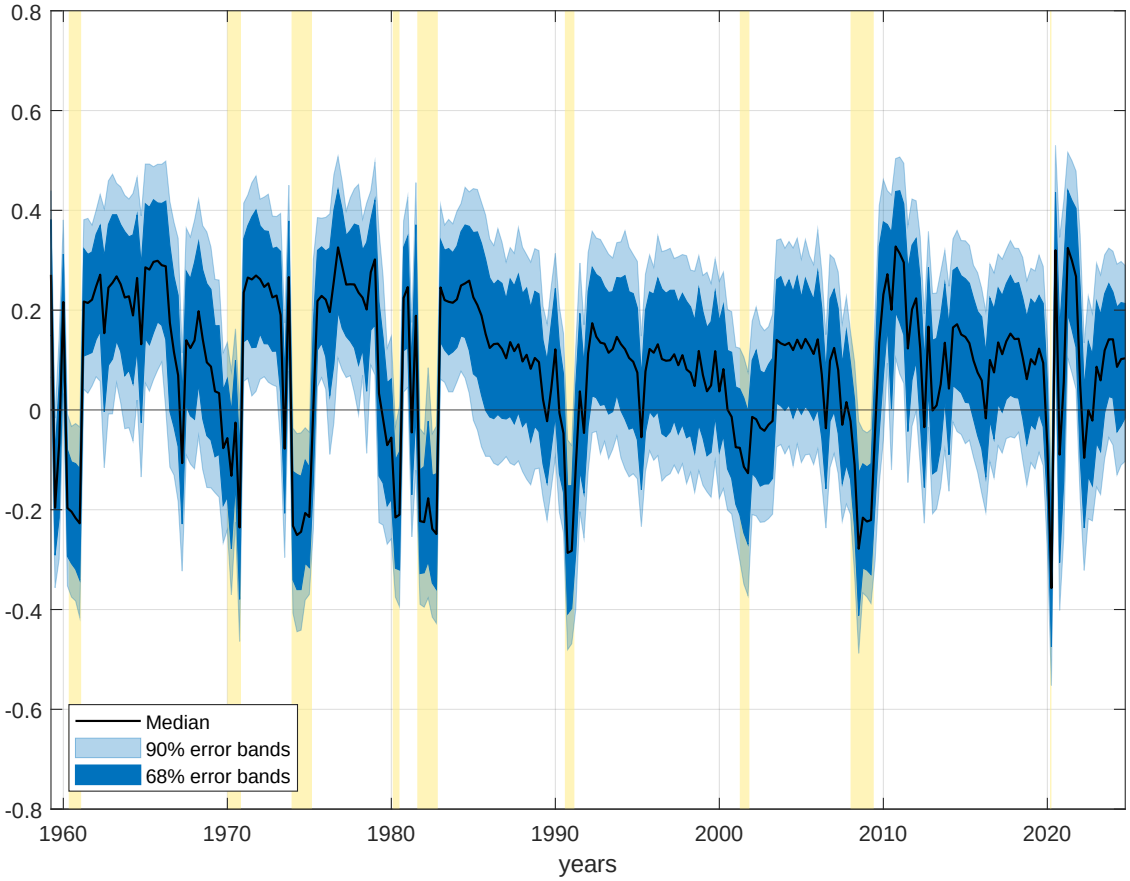
$$\text{Skewness}_{K(0.10)} = \frac{Q_{0.90} + Q_{0.10} - 2Q_{0.50}}{Q_{0.90} - Q_{0.10}}$$

- **Kelley(95/5) Skewness**, using the 95th and 5th percentiles:

$$\text{Skewness}_{K(0.05)} = \frac{Q_{0.95} + Q_{0.05} - 2Q_{0.50}}{Q_{0.95} - Q_{0.05}}$$

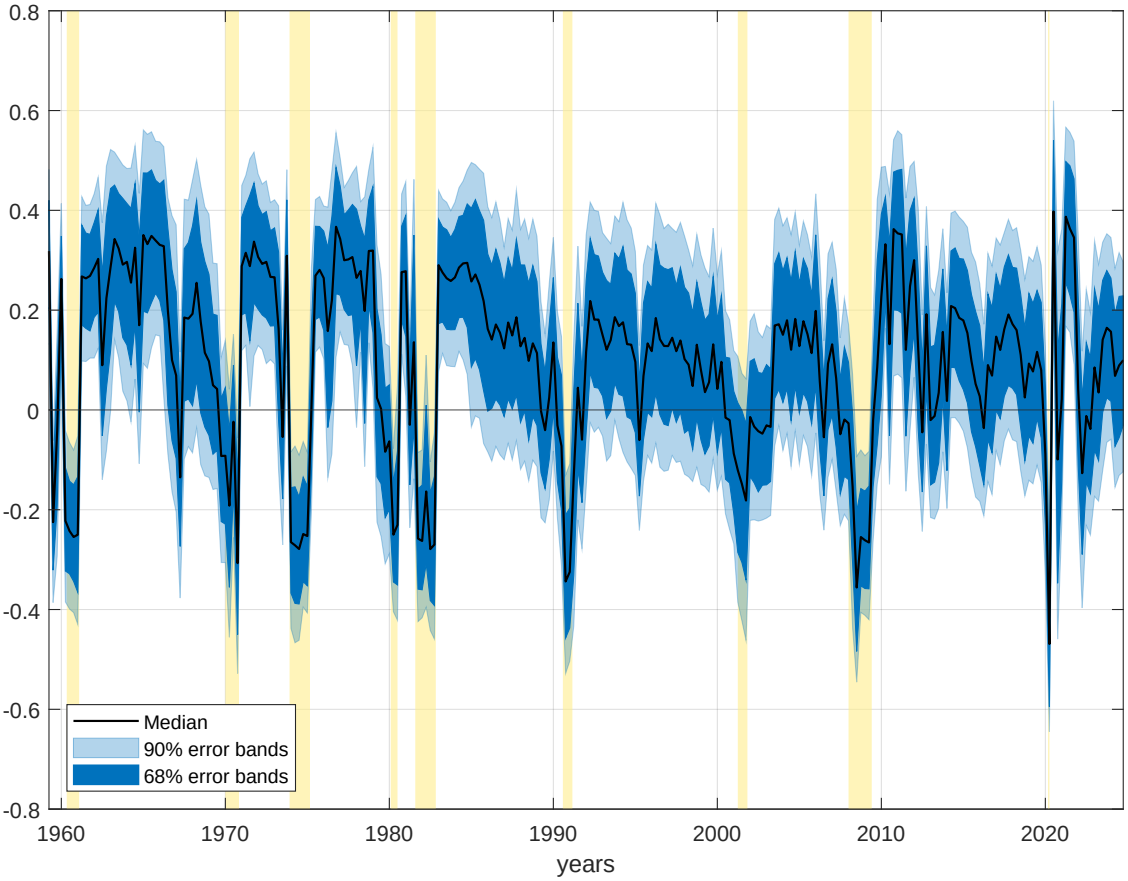
These alternative measures provide a quantile-based view of asymmetry that complements the moment-based skewness used in the main text. The results are consistent with the overall conclusions and are available in Figures I.1 and I.2.

Figure I.1: Macroeconomic Skewness — Kelley(90/10)



Note: Sample period: 1959.Q2 — 2024.Q4. Evolution of macroeconomic skewness produced using the Kelley's measure. The median is reported in black solid line and the 68% and 90% error bands in blue areas. The yellow areas denote the NBER-dated recessions.

Figure I.2: Macroeconomic Skewness — Kelley(95/5)



Note: Sample period: 1959.Q2 — 2024.Q4. Evolution of macroeconomic skewness produced using the Kelley's measure. The median is reported in black solid line and the 68% and 90% error bands in blue areas. The yellow areas denote the NBER-dated recessions.

J Forecasting

J.1 Methodology

Consider the following state-space representation of a skewed dynamic factor model with regime switching:

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{H}_{s_t}\mathbf{x}_t + \boldsymbol{\varepsilon}_t, \qquad t = 1, \ldots, T, \\
\mathbf{x}_t &= \mathbf{F}_{s_t}\mathbf{x}_{t-1} + \boldsymbol{\eta}_t, \\
\boldsymbol{\varepsilon}_t &\sim \text{normal}\left(\boldsymbol{\mu}_{s_t,\varepsilon}, \boldsymbol{\Sigma}_{s_t,\varepsilon}\right), \\
\boldsymbol{\eta}_t &\sim \text{closed skew-normal}\left(\boldsymbol{\mu}_{s_t,\eta}, \boldsymbol{\Sigma}_{s_t,\eta}, \boldsymbol{\Gamma}_{s_t,\eta}, \boldsymbol{\nu}_{s_t,\eta}, \boldsymbol{\Delta}_{s_t,\eta}\right), \\
p_{i,j} &= \Pr(s_t = j | s_{t-1} = i), \qquad i, j = 1, \ldots, K,
\end{aligned}
$$

where $\mathbf{y}_t$ is an $N \times 1$ vector of observed variables, $\mathbf{x}_t$ is an $J \times 1$ vector of state variables, $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ are vectors of measurement and transition shocks, containing ς_{it} and ϑ_t. For simplicity of exposition, it is assumed that there is no measurement errors ($\boldsymbol{\varepsilon}_t = \mathbf{0}$).

Model forecasts can be computed based on draws from the posterior predictive distribution of $\mathbf{y}_{T+1:T+H}$, with the sequence $t_1 : t_2$ indicating the periods from t_1 to t_2. I use the parameter draws $\left\{\theta^{(j)}\right\}_{j=1}^{nsim}$ generated with the RWMH algorithm as a starting point. Specifically, I generate draws from the posterior predictive density using the following decomposition:

$$
\begin{aligned}
p(\mathbf{y}_{T+1:T+h}|\mathbf{y}_{1:T}) = \int_{(\theta,\mathbf{x}_T,s_T)} \Bigg[& \int_{\mathbf{x}_{T+1:T+h},S_{T+1:T+h}} p(\mathbf{y}_{T+1:T+h}|\mathbf{x}_{T+1:T+h}) \\
& \times p(\mathbf{x}_{T+1:T+h}, s_{T+1:T+h}|\theta, \mathbf{x}_T, s_T, \mathbf{y}_{1:T}) d(\mathbf{x}_{T+1:T+h}, s_{T+1:T+h}) \Bigg] \\
& \times p(\theta, \mathbf{x}_T, s_T|\mathbf{y}_{1:T}) d(\theta, \mathbf{x}_T, s_T).
\end{aligned}
$$

The decomposition shows how the predictive density reflects uncertainty about parameters

and states at the forecast origin, $p(\theta, \mathbf{x}_T, s_T|\mathbf{y}_{1:T})$, and uncertainty about future states. Motivated by this decomposition, I generate draws from the predictive density taking into account the hidden Markov regimes s_t.

Algorithm (Predictive Density Draws)

For $j = 1, 2, \ldots, n_{sim}$,

1. Draw $\theta^{(j)}, \mathbf{x}_T^{(j)}, s_T^{(j)}$ from the posterior distribution $p(\theta, \mathbf{x}_T, s_T|\mathbf{y}_{1:T})$.

2. Draw $p\left(\mathbf{x}_{T+1:T+h}, s_{T+1:T+h}|\theta^{(j)}, \mathbf{x}_T^{(j)}, s_T^{(j)}\right)$ as follows:

   (a) Draw $p\left(s_{T+1:T+h}|\theta^{(j)}, s_T^{(j)}\right)$ from the transition matrix $\boldsymbol{P}^{(j)}$ and the regime $s_T^{(j)}$;

   (b) Draw the sequence of shock innovations $\boldsymbol{\eta}_{T+1:T+h}^{(j)}$:

$$\boldsymbol{\eta}_{T+1:T+h}^{(j)} \sim \text{closed skew-normal}\left(\boldsymbol{\mu}_{s_t^{(j)},\eta}, \boldsymbol{\Sigma}_{s_t^{(j)},\eta}, \boldsymbol{\Gamma}_{s_t^{(j)},\eta}, \boldsymbol{\nu}_{s_t^{(j)},\eta}, \boldsymbol{\Delta}_{s_t^{(j)},\eta}\right)$$

   with $t = T+1, \ldots, T+H$.

   (c) Starting from $\mathbf{x}_T^{(j)}$, iterate the state transition equation forward:

$$\mathbf{x}_t^{(j)} = \mathbf{F}\left(\theta^{(j)}\right)_{s_t^{(j)}} \mathbf{x}_{t-1}^{(j)} + \boldsymbol{\eta}_t^{(j)}, \qquad t = T+1, \ldots, T+H.$$

3. Compute the sequence $\mathbf{y}_{T+1:T+h}^{(j)}$ using the measurement equation:

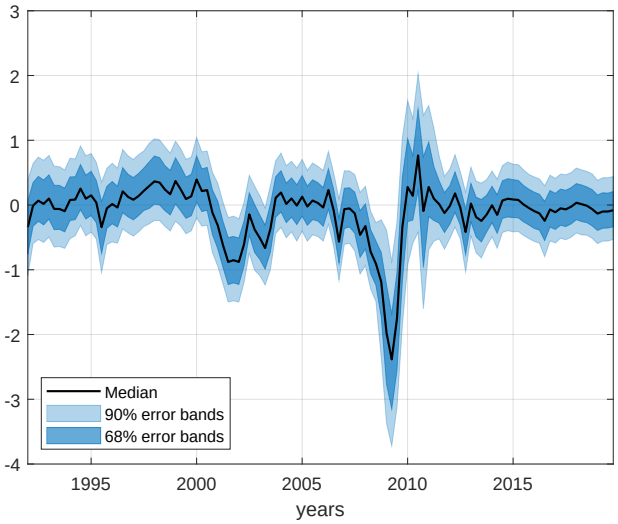
$$\mathbf{y}_t^{(j)} = \mathbf{H}\left(\theta^{(j)}\right)_{s_t} \mathbf{x}_t^{(j)}, \qquad t = T+1, \ldots, T+H.$$

This algorithm produces n_{sim} trajectories $\mathbf{y}_{T+1:T+h}^{(j)}$ from the predictive distribution of $\mathbf{y}_{T+1:T+h}^{(j)}$ given $\mathbf{y}_T$. In my subsequent empirical work, I take 10,000 draws from the posterior distribution $p(\theta, \mathbf{x}_T, s_T|\mathbf{y}_{1:T})$. I discard the first 1,000 draws and select every 10th draws to get 1,000 draw of parameters and initial states. For each of these draws, I execute Steps 2 and 3 of the algorithm 10 times, which produces a total of $10,000$ draws from the predictive distribution.

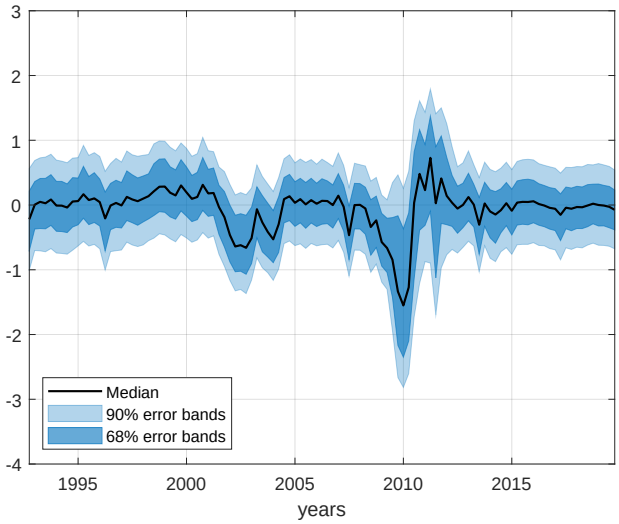
J.2 Additional Figures

Figure J.1: Out-of-sample Predictions — Macroeconomic Factor.

(a) One-quarter ahead



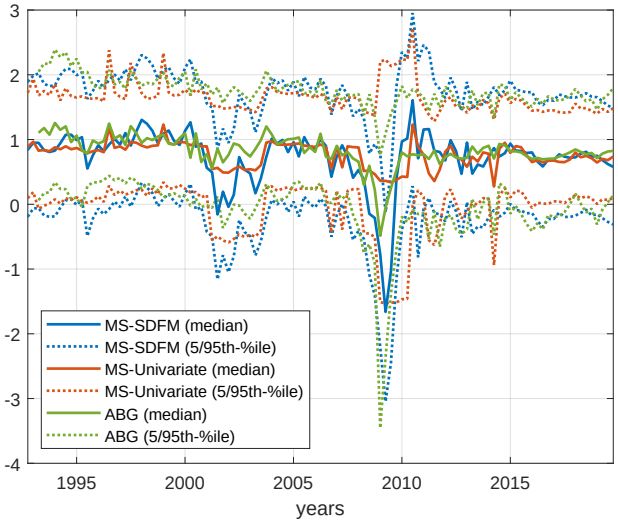
(b) Four-quarter ahead



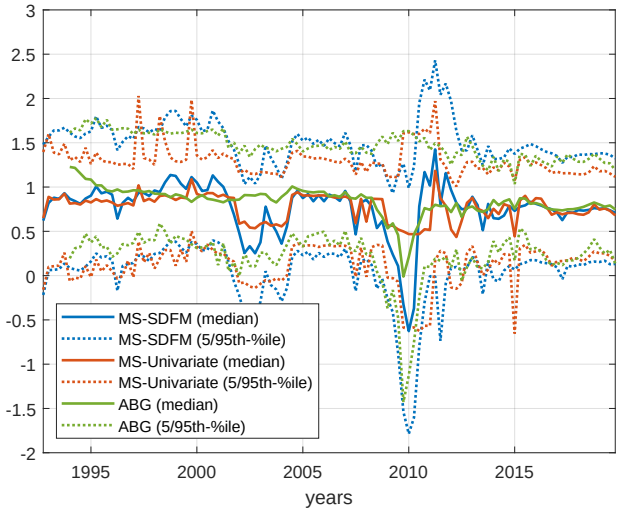
Note: Sample period: 1992.Q1 — 2019.Q4. Time series evolution of the predicted distribution of one quarter ahead (panel a) and four quarter ahead (panel b) macroeconomic factor. The median is reported in black solid line and the 68% and 90% error bands in blue areas.

Figure J.2: Comparison of Out-of-sample Predictions — Real GDP Growth.

(a) One-quarter ahead



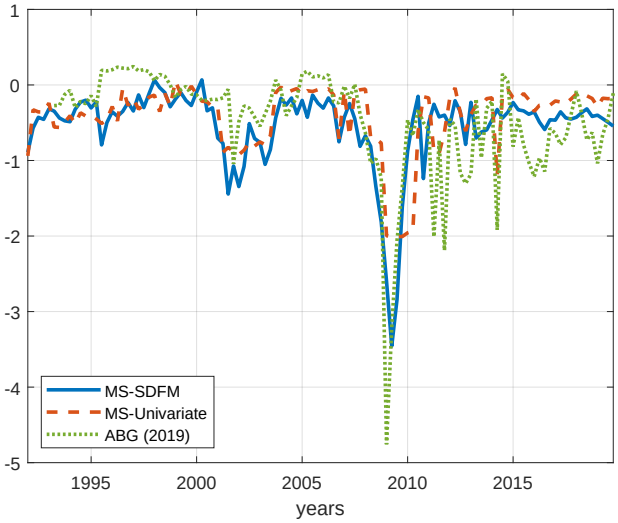
(b) Four-quarter ahead



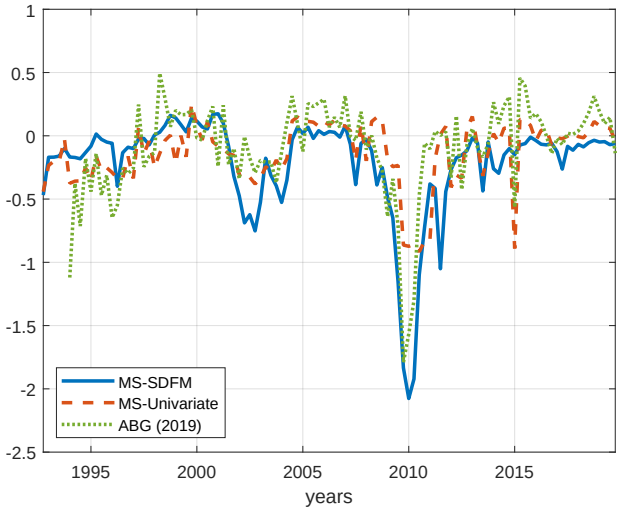
Note: Sample period: 1992.Q1 — 2019.Q4. Time series evolution of the predicted distribution of one quarter ahead (panel a) and four quarter ahead (panel b) real GDP growth.

Figure J.3: Comparison of Expected Shortfalls — Real GDP Growth.

(a) One-quarter ahead



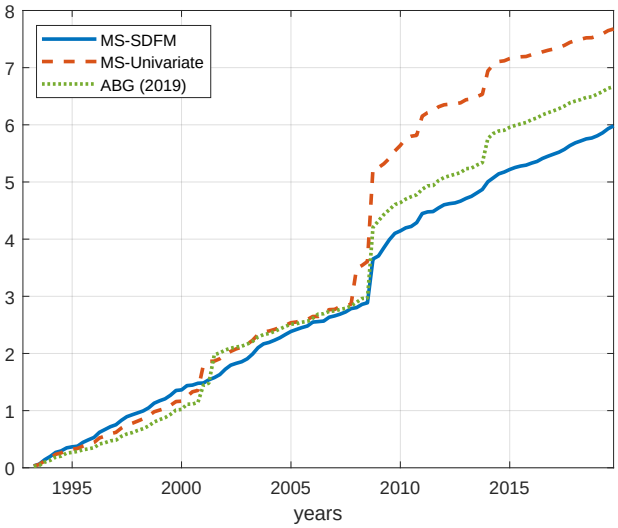
(b) Four-quarter ahead



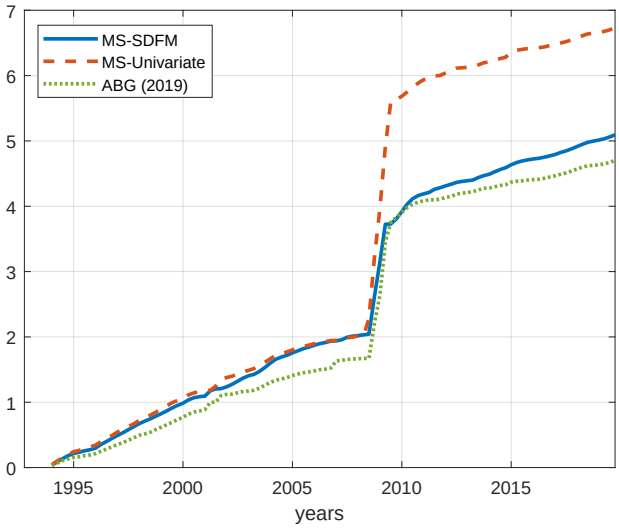
Note: Sample period: 1992.Q1 — 2019.Q4. Time series evolution of the predicted distribution of one quarter ahead (panel a) and four quarter ahead (panel b) real GDP growth.

Figure J.4: Cumulative Quantile Scores — Real GDP Growth.

(a) One-quarter ahead

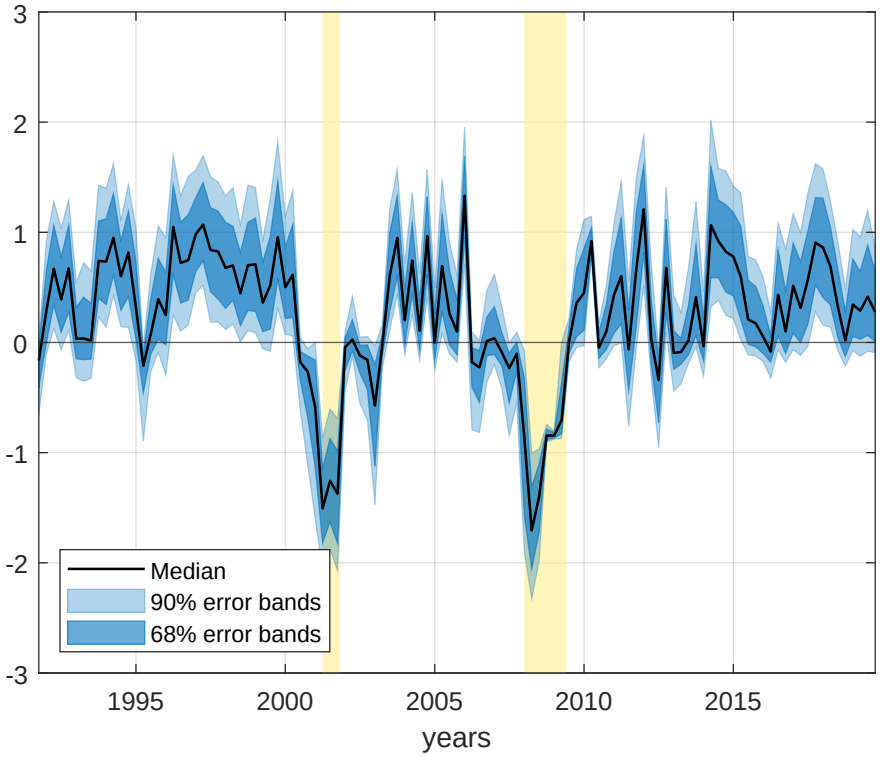


(b) Four-quarter ahead



Note: Sample period: 1992.Q1 — 2019.Q4. Cumulative sums of quantile scores for one quarter ahead (panel a) and four quarter ahead (panel b) of the 5th percentile of real GDP growth. Lower values signify better performance.

Figure J.5: Macroeconomic Skewness — Out-of-sample Estimates



Note: Sample period: 1992.Q1 — 2019.Q4. Evolution of macroeconomic skewness produced using the parameter estimates obtained recursively from 1992.Q1 to 2019.Q4. The median is reported in black solid line and the 68% and 90% error bands in blue areas. The yellow areas denote the NBER-dated recessions.

References

Azzalini, A. (1985): "A Class of Distributions Which Includes the Normal Ones," *Scandinavian Journal of Statistics*, 12(2), 171–178.

——— (2013): *The Skew-Normal and Related Families*, Institute of Mathematical Statistics Monographs. Cambridge University Press.

Guerrón-Quintana, P., A. Khazanov, and M. Zhong (2024): "Nonlinear Dynamic Factor Models," mimeo.

Kelley, T. (1947): *Fundamentals of Statistics*, no. n° 2 in Fundamentals of Statistics. Harvard University Press.

Lhuissier, S. (2022): "Financial Conditions and Macroeconomic Downside Risks in the Euro Area," *European Economic Review*, 143(C).

Perez-Quiros, G., and A. Timmermann (2001): "Business cycle asymmetries in stock returns: Evidence from higher order moments and conditional densities," *Journal of Econometrics*, 103(1-2), 259–306.

Timmermann, A. (2000): "Moments of Markov switching models," *Journal of Econometrics*, 96(1), 75–111.

Watanabe, S. (2010): "Asymptotic Equivalance of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory," *Journal of Machine Learning Research*, 11, 3571–3594.