WORKING paper

BANQUE DE FRANCE
EUROSYSTÈME

Financial Conditions and Macroeconomic Downside Risks in the Euro Area

Stéphane Lhuissier[1]

January 2022, WP #863

ABSTRACT

Motivated by empirically characterizing the relationship between financial conditions and downside macroeconomic risks in the euro area, I develop a regime-switching skew-normal model with time-varying probabilities of transitions. Using Bayesian methods, the model estimates show that a strong cyclical pattern emerges from the conditional skewness (a measure of the asymmetry of the predictive distribution), which has a tendency to rapidly decline to negative territory prior and during recessions. However, the inclusion of financial-specific information in time-varying probabilities does not help to anticipate such skewness nor more generally to provide advance warnings of tail risks.

Keywords: Financial Conditions, Downside Risks, Predictability, Regime-Switching Models.

JEL classification: C11; C2; E32

[1] 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>.
I thank Tobias Adrian, Isaac Baley, Fabrice Collard (Banque de France discussant), Marco Del Negro, Kyle Jurado, Florens Odendahl, Pierre-Alain Pionnier (PSE Discussant), Jean-Marc Robin, Moritz Schularick, Mathias Trabandt and participants at the 10th French Econometrics Conference (PSE), CFE 2018 (University of Pisa), the 2019 Banque de France seminar, ICMAIF 2019 (University of Crete), and the 2019 Padova Workshop (Padova University) for their helpful comments.

NON-TECHNICAL SUMMARY

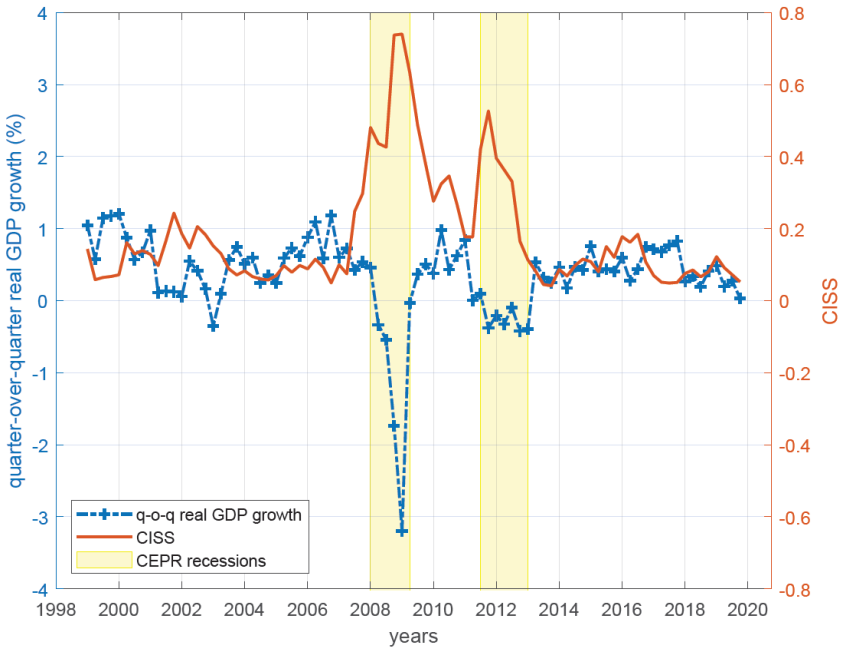
Following the Great Recession, macroeconomic risks have received substantial attention from the profession. For example, the European Central Bank (ECB), like many central banks, communicates regularly in press conferences its risk perception on inflation and growth over the foreseeable future. The ECB's Governing Council informs whether it sees these risks as balanced, on the upside or on the downside. From an academic perspective, there is evidence of a negative relationship between financial conditions and the lower quantiles of real economic growth, suggesting that financial conditions have significant predictive content for downside risks to growth.

In this paper, I introduce and develop a regime-switching model with time-varying transition probabilities to estimate the moments of the distribution of euro area real gross domestic product (GDP) growth conditional on economic and financial conditions, and to characterize business cycle variation in the probability distribution and time-varying risks around GDP growth. In particular, I focus on a non-Gaussian model based on the skew-normal distribution, and in which location, scale and shape parameters are allowed to vary over time according to independent two-state Markov-switching processes. In the setup, the probability of moving between states is allowed to depend on economic and financial conditions. By doing so, I am able to explore the predictive content of a tightening of financial conditions for changes in the distribution of GDP growth over time. The method provides an empirical characterization of business cycle variation of risks that goes beyond what can be achieved through Gaussian models, and allows for the construction of informative measures of downside macroeconomic risks.

By allowing asymmetric distributions across time as well as mixing probabilities, the model is able to reproduce well-documented evidence on time variation in the moments of the conditional (predictive) distribution of GDP growth, but also to fully characterize the cyclical variation in risk, as measured for example by the conditional skewness. Model estimates capture recurrent features of business cycles in modern economies like periodic shifts in the long-term mean rate of economic growth and the rise in volatility observed during economic downturns. Furthermore, I provide new evidence that, during economic downturns, the skewness of GDP growth tends to be negative, implying that extreme values on the left side of the mean are more likely than the extreme values of the same magnitude on the right side of the mean, leading to a left tail of a large negative output. The skewness also tends to decline in anticipations of recessions, a feature that is also observable for the U.S. economy.

However, the inclusion of financial-specific information in time-varying probabilities helps poorly to predict different features of the GDP growth distribution, including conditional skewness. Indeed, the model estimates rule out any role of financial conditions in the transition probabilities governing location, scale and shape parameters as the mass of the posterior distributions of elasticities linking financial conditions to transition probabilities lie both on the positive and negative side. At the same time, the signals coming from financial conditions are sporadic, and in particular, given the limited number of transitions observed in the sample, it turns out difficult to obtain precise estimates of some of the parameters. Thus, when assessed its statistical performance, the model that incorporates an aggregate indicator of financial stress does not improve out-of-sample forecasts in terms of point, density and tail risks. Therefore, financial conditions cannot be seen as a warning signal of downside risks in GDP growth. A caveat of the analysis is the use of quarterly data, while financial variables are often available at higher frequency. Considering a real-time nowcasting exercise with high-frequency data might be more accurate to predict downside risks.

Economic and Financial Conditions in the Euro Area



Note: The quarter-over-quarter real GDP growth rate (dotted blue line) is labeled on the left. The Composite Indicator of Systemic Stress (CISS) index (solid red line) is labeled on the right. CISS includes 15 raw, mainly market-based financial stress measures that are split equally into five categories: the financial intermediaries sector, money markets, equity markets, bond markets and foreign exchange markets.

Conditions financières et risques de baisse de l'activité économique dans la zone euro

RÉSUMÉ

Cette étude développe un modèle univarié asymétrique à changements de régimes markoviens avec des probabilités de transition variables dans le temps afin de caractériser empiriquement la relation entre les conditions financières et les risques de baisse de l'activité économique dans la zone euro. En utilisant des méthodes bayésiennes, l'estimation du modèle montre que la skewness (une mesure de l'asymétrie de la distribution prédictive) affiche un profil cyclique, elle descend rapidement en territoire négatif avant et durant les récessions. Cependant, l'intégration d'informations financières dans les probabilités de transition n'aide pas à anticiper une telle skewness, ni de manière plus générale, à signaler à l'avance les risques extrêmes.

Mots-clés : conditions financières, risques de baisse, modèle à changements de régimes

I. Introduction

Following the Great Recession, macroeconomic risks have received substantial attention from the profession. For example, the European Central Bank (ECB), like many central banks, communicates regularly in press conferences its risk perception on inflation and growth over the foreseeable future. The ECB's Governing Council informs whether it sees these risks as balanced, on the upside or on the downside. From an academic perspective, the recent seminal work by Adrian, Boyarchenko, and Giannone (2019) on U.S. data provides evidence of a negative relationship between financial conditions and the lower quantiles of real economic growth, suggesting that financial conditions have significant predictive content for downside risks to growth.

In this paper, I introduce and develop a regime-switching model with time-varying transition probabilities to estimate the moments of the distribution of euro area real gross domestic product (GDP) growth conditional on economic and financial conditions, and to characterize business cycle variation in the probability distribution and time-varying risks around GDP growth. In particular, I focus on a non-Gaussian model based on the skew-normal distribution developed by Azzalini (1985, 1986), and in which location, scale and shape parameters are allowed to vary over time according to independent two-state Markov-switching processes. In the setup, the probability of moving between states is allowed to depend on economic and financial conditions. By doing so, I am able to explore the predictive content of a tightening of financial conditions for changes in the distribution of GDP growth over time. The method provides an empirical characterization of business cycle variation of risks that goes beyond what can be achieved through Gaussian models, and allows for the construction of informative measures of downside macroeconomic risks.

By allowing asymmetric distributions across time as well as mixing probabilities, the model is able to reproduce well-documented evidence on time variation in the moments of the conditional (predictive) distribution of GDP growth, but also to fully characterize the cyclical variation in risk, as measured for example by the conditional skewness. Model estimates capture recurrent features of business cycles in modern economies like periodic shifts in the long-term mean rate of economic growth (e.g., Hamilton, 1989) and the rise in volatility observed during economic downturns (e.g., Justiniano and Primiceri, 2008; Liu, Waggoner, and Zha, 2011; Jurado, Ludvigson, and Ng, 2015; Lhuissier, 2017, 2018). Furthermore, I provide new evidence that, during economic downturns, the skewness of GDP growth tends to be negative, implying that extreme values on the left side of the mean are more likely than the extreme values of the same magnitude on the right side of the mean, leading to a left tail of a large negative output. The skewness also tends to decline in anticipations of recessions,

a feature that is also observable for the U.S. economy (e.g., De Polis, Delle Monache, and Petrella, 2020). Clearly, the skewness is procyclical. The derivation of moments for a range of univariate Markov-switching skew-normal model is not straightforward and I extend the methodology developed by Timmermann (2000), which initially works with Gaussian mixture models. The skew-normal feature employed in this paper enables me to generate wider range of coefficients of skewness and kurtosis, and therefore to better capture nonlinearities in the entire conditional distribution of real GDP growth in the euro area with respect to the standard Markov-switching Gaussian model.

However, the inclusion of financial-specific information in time-varying probabilities helps poorly to predict different features of the GDP growth distribution, including conditional skewness. Indeed, the model estimates rule out any role of financial conditions in the transition probabilities governing location, scale and shape parameters as the mass of the posterior distributions of elasticities linking financial conditions to transition probabilities lie both on the positive and negative side. At the same time, the signals coming from financial conditions are sporadic, and in particular, given the limited number of transitions observed in the sample, it turns out difficult to obtain precise estimates of some of the parameters. Thus, when assessed its statistical performance, the model that incorporates an aggregate indicator of financial stress does not improve out-of-sample forecasts in terms of point, density and tail risks. Therefore, financial conditions cannot be seen as a warning signal of downside risks in GDP growth.

I rely on Bayesian methods to estimate the model. More specifically, I develop a Gibbs sampler for Bayesian inference of time series model subject to Markov shifts in location, scale and shape parameters. The Gibbs sampling procedure can thus be seen as an extension of Albert and Chib (1993) by allowing asymmetric shifts in the conditional distribution. Specifically, I take advantage of the stochastic representation of skew-normal variables, which is based on a convolution of normal and truncated-normal variables, in order to obtain a straightforward Markov Chain Monte Carlo (MCMC) sampling sequence that involves a 6-block Gibbs sampler for Markov-switching models, and in which one can generate in a flexible and straightforward manner alternatively draws from full conditional posterior distributions. By doing so, I am able to generate draws from the posterior distributions of functions of parameters in order to fully characterize the statistical uncertainty surrounding the estimation of the higher order moments of the growth distribution. I believe that the Gibbs sampler developed in this paper is a promising tool to infer time variation in the conditional distributions of any macroeconomic time series data.

Relation to other studies. This paper is related to an increasing literature that examines the relationship between financial conditions and economic activity (e.g., Gilchrist

and Zakrajšek, 2012; Gertler and Gilchrist, 2018; Brunnermeier, Palia, Sastry, and Sims, forthcoming). Adrian, Boyarchenko, and Giannone (2019) suggest that financial conditions have significant predictive content for downside risks to growth. Following a similar approach, Adrian, Grinberg, Liang, and Malik (2018) examine the distribution of expected GDP growth for 11 advanced economies and find similar results. By contrast, Plagborg-Møller, Reichlin, Ricco, and Hasenzagl (2020) report that financial variables have very limited predictive power for the U.S. economy. Finally, Figueres and Jarociński (2020) examine the informative content of a certain numbers of financial conditions about the tail risks to output growth in the euro area.

The major difference between my approach and the one adopted by Adrian, Boyarchenko, and Giannone (2019) lies in the methodology itself. While the authors use a two-step quantile approach that requires to fit a skew-t distribution to the estimated quantiles, I directly estimate the time-varying parameters of a skew-normal distribution within a Markov-switching framework, which allows great flexibility and tractability in modeling time variation in downside risks as a function of economic and financial conditions. Looking at U.S. data, Plagborg-Møller, Reichlin, Ricco, and Hasenzagl (2020) and De Polis, Delle Monache, and Petrella (2020) introduce and estimate versatile parametric skew-t densities with time-varying parameters as linear functions of financial predictors and past GDP growth, which favor smooth variation in parameters of the distribution, and thus in conditional moments. By contrast, by allowing the mixing probabilities to display time dependence, my approach is able to generate both smooth and abrupt variations of conditional moments, a key feature of business cycle fluctuations. For example, financial crises are well-known for hitting the economy instantaneously, which favors models with abrupt changes like Markov-switching models. Using a semi-structural model subject to Markov mean and variance shifts, Caldara, Cascaldi-Garcia, Cuba-Borda, and Loria (2020) investigate the role of the financial and real conditions to predict tail risks in the U.S. economy.

From a methodological point of view, I propose a Markov-switching model with time-varying location, scale and shape parameters, and mixing weights to forecast the entire distribution of GDP growth. In the literature, time-varying asymmetry has been, in the first place, modelled through the generalized autoregressive conditional heteroskedasticity (GARCH) models. Notable examples include Harvey and Siddique (1999), Jondeau and Rockinger (2003), and Christoffersen, Heston, and Jacobs (2006). The deterministic behavior of such systems lead, however, to limited implications. Feunou and Tédongap (2012) and Iseringhausen (2018) go a step further by modelling time-varying skewness as stochastic by extending the standard stochastic volatility model. Nakajima (2013) introduces a Markov-switching framework as an alternative to modelling stochastically asymmetry. While my

specification remains nested into his framework, I am able to capture a wider range of coefficients of skewness since, not only the shape parameter is allowed to vary over time, but also the location and scale parameters of the distribution, of which the times of changes are stochastically independent. Most importantly, unlike Nakajima (2013)'s approach, I incorporate time-varying transition probabilities of transitions between regimes to assess the potential role of financial conditions in helping to predict time variation in moments of the GDP growth distribution.

The paper is organised as follows. Section II presents some motivating evidence. Section III outlines the general methodology employed in this paper. Section IV presents the main results. Section V analyzes the time series variation in GDP growth distribution. Section VI discusses out-of-sample results. Section VII conducts robustness checks. Section VIII concludes.

II. Motivating Evidence

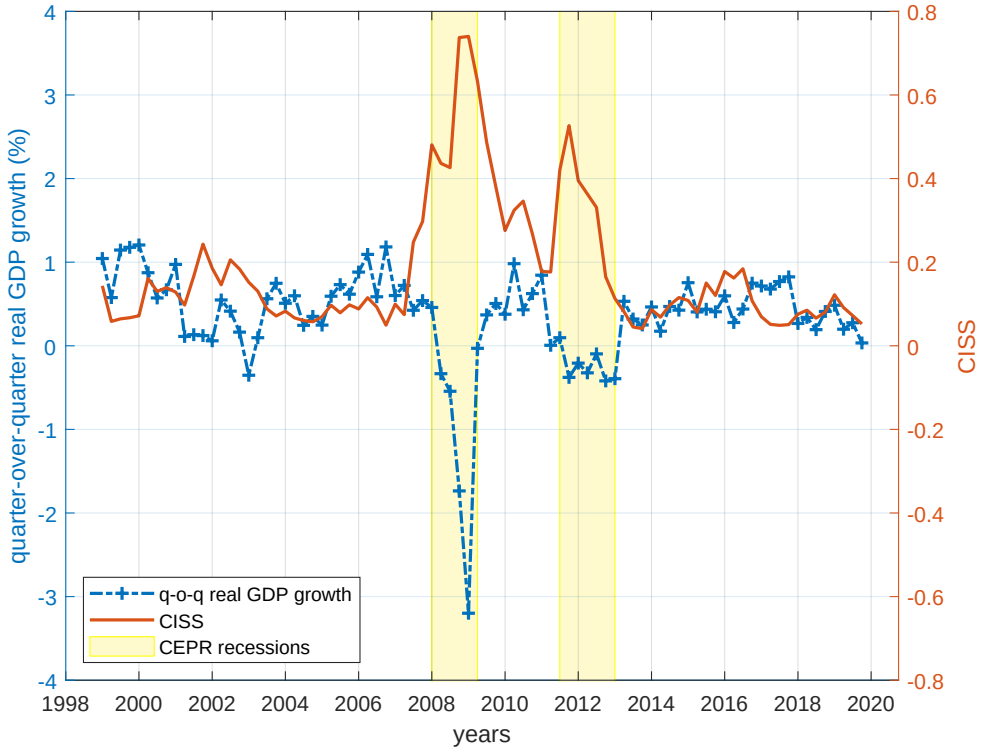
In this section I present some evidence that motivates the analysis of the paper.

Figure 1 plots the real GDP growth rate of the euro area, along with the Composite Indicator of Systemic Stress (CISS) developed by Kremer, Lo Duca, and Holló (2012) from the first quarter of 1999 to the fourth quarter of 2019. The CISS is a weekly index maintained by the ECB. It includes 15 raw, mainly market-based financial stress measures that are split equally into five categories: the financial intermediaries sector, money markets, equity markets, bond markets and foreign exchange markets.[1]

The connection of the financial variable to the real variable seems very unstable. Indeed, during financial stress events (i.e., the global financial crisis and the European sovereign debt) there is a close connection between sharp reductions in output and violent systemic risk. By contrast, other periods show that movements in financial stress have very little effect on the economy. This nonlinear relationship between financial sector and real economy has been the inspiration for macro-finance theorists to depart from log-linearized models (e.g., Bernanke and Gertler, 1989; Kiyotaki and Moore, 1997; Bernanke, Gertler, and Gilchrist, 1999) to study global dynamics of the system characterized by "normal" and "crisis" states, which are determined by financial constraints in the intermediation sector (e.g., He and Krishnamurthy, 2012, 2013; Adrian and Boyarchenko, 2012; Brunnermeier and Sannikov, 2014; Maggiori, 2017). From an empirical perspective, Hubrich and Tetlow (2015) and Lhuissier (2017) employ Markov-switching structural vector autoregressions to better understand how

[1] On the ECB - statistical Data Warehouse, the code series for the real GDP series and CISS index are MNA.Q.Y.I8.W2.S1.S1.B.B1GQ._Z._Z._Z.EUR.LR.N and CISS.D.U2.Z0Z.4F.EC.SS_CI.IDX, respectively. I perform a first log difference transformation to GDP series to compute the growth rates.

FIGURE 1. Output growth and CISS.



Note: Sample period: 1999.Q1 — 2019.Q4. The quarter-over-quarter real GDP growth rate (dotted blue line) is labeled on the left. The CISS index (solid red line) is labeled on the right. The yellow areas denote the CEPR recessions.

disruptions in financial intermediation sector manifest themselves and what their effects are on the rest of the economy.[2]

This first fact motivates an analysis which is based on the use of nonlinear technique to better understanding the relationship between the financial sector and the real economy.

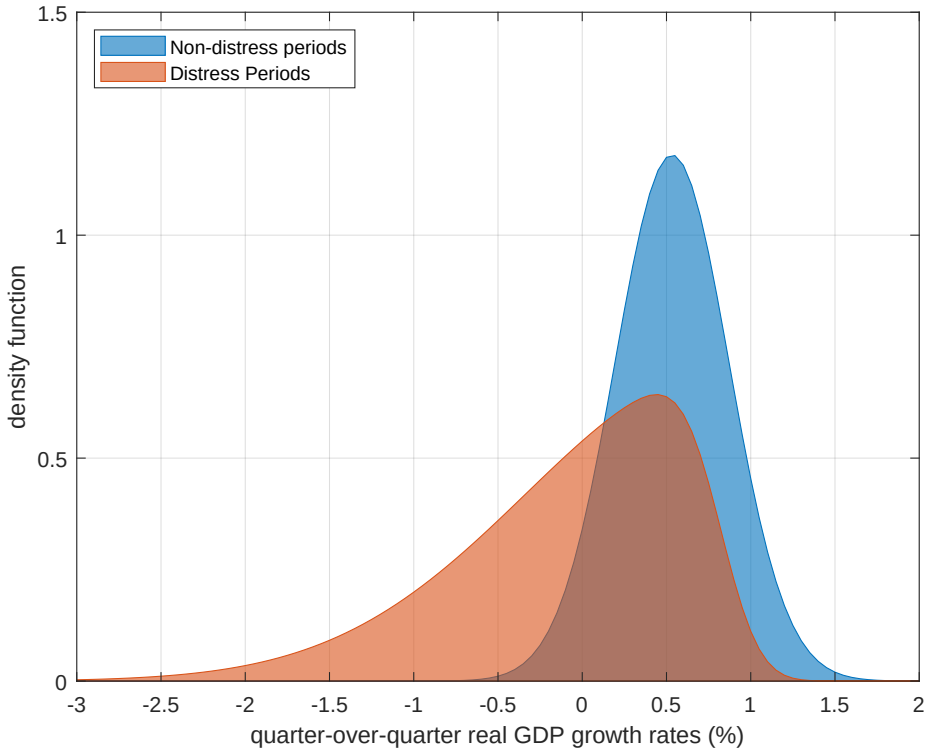
Figure 2 shows the fitted distribution[3] of real GDP growth in the euro area conditional on its location in the "distress" or in the "non-distress" period. I define distress periods as the highest one-third of realizations of the CISS index and I require that the distress or non-distress periods minimally cover two quarterly periods. I classify two distress periods: the third quarter of 2001 to the fourth quarter of 2002 and the third quarter of 2007 to the fourth quarter of 2012. Clearly, the last financial distress period is associated with the

[2]Using a Markov-switching business cycle model, Lhuissier and Tripier (2021) point to worsening credit-market conditions during distress periods.

[3]I fit a skew-normal distribution, introduced by Azzalini (1985, 1986), to the quarter-over-quarter real GDP growth rates in the euro area.

recent global financial crisis and the sovereign debt crisis, which typically followed in several European countries. Interestingly, the first financial distress period prevailed during periods marked by the 9/11 terrorist attacks, Dot-com bubble, and corporate scandals.

FIGURE 2. Density function of real GDP growth.



Note: The density functions are computed by fitting the skew-normal distribution to the real GDP growth rates of the euro area. "Distress periods" are defined as the highest one-third realizations of the CISS.

As can be seen there is a remarkable difference in the shape of the density function across the distress and non-distress periods. While in non-distress periods, GDP growth appears to be Gaussian and lying almost exclusively within the positive region, it exhibits smaller expected values, larger variance and negative skewness in distress periods. Interestingly, one is more likely to see GDP falling than rising during distress periods, this is in any case certain; in fact, distress periods are associated with positive GDP growth as well. By contrast, there is a very small probability of observing a negative growth rate during what is considered to be non-distress periods.

This second fact suggests an analysis which is based on an empirical characterization of variations in risk to growth that goes beyond what can be achieved through the typical Gaussian assumption.

Finally, it is important to note that these simple descriptive statistics do not allow us to know the role of financial conditions in shaping the GDP distribution. The objective of the next sections is to better understand the predictive power of financial conditions for moments that go well beyond the mean.

III. The Methodology

This section presents the general methodology employed in this paper. Section III.1 discusses the baseline statistical model, while Section III.2 describes the Gibbs-sampler procedure that will be used for Bayesian inference.

III.1. A Markov-switching skew-normal model.

I employ a statistical model in which the observation at time t, y_t, is generated as follows:

$$p(y_t|Y_{t-1}, z_t, s_t, \theta) = \text{skew-normal}(y_t|\mu_{s_t^{\text{location}}}, \sigma^2_{s_t^{\text{scale}}}, \alpha_{s_t^{\text{shape}}}), \qquad t = 1, \dots, T, \tag{1}$$

where $Y_t = [y_1, \dots, y_t]$ contains the information set available at time $t-1$, θ is a vector of parameters, z_t is a vector of lagged conditioning variables, $s_t = \left\{s_t^{\text{location}}, s_t^{\text{scale}}, s_t^{\text{shape}}\right\}$ contains the variables that follow Markov processes, skew-normal$(Y|\xi, \omega^2, \alpha)$ denotes the skew-normal distribution of Y with location parameter ξ, scale parameter ω^2, and shape parameter α, and T is the sample size.

The skew-normal family was introduced by Azzalini (1985, 1986) as the extension of the normal family from a symmetric form to an asymmetric form. It is a distribution that has an additional parameter: a shape parameter $\alpha \in \mathbb{R}$, which allow for possible deviation from symmetry. Formally, the distribution of Y is as follows:

$$p(Y|\xi, \omega^2, \alpha) = \frac{2}{\omega}\phi\left(\frac{Y-\xi}{\omega}\right)\Phi\left(\alpha\frac{Y-\xi}{\omega}\right), \tag{2}$$

where $\phi(.)$ and $\Phi(.)$ denote the standard normal density function and cumulative distribution function, respectively. If the shape parameter is equal to zero, then the density of Y is a normal distribution with mean ξ, and standard deviation σ.

The discrete and unobserved variable s_t^k that governs the parameter $k \in \{\text{location}, \text{scale}, \text{shape}\}$ is an exogenous two-states first-order Markov process with the following time-varying transition matrix Q_t^k

$$Q_t^k = \begin{bmatrix} q_{1,1}^k(z_t) & q_{1,2}^k(z_t) \\ q_{2,1}^k(z_t) & q_{2,2}^k(z_t) \end{bmatrix}, \tag{3}$$

where $q_{i,j}^k(z_t) = \Pr(s_t^k = i|s_{t-1}^k = j, z_t)$ denote the transition probabilities that s_t^k is equal to i given that s_{t-1}^k is equal to j, with $i, j \in \{1, 2\}$, $q_{i,j}^k \geq 0$ and $\sum_{j=1}^2 q_{i,j}^k = 1$. With the

introduction of three-independent Markov-switching processes, the overall transition matrix Q_t becomes

$$Q_t = Q_t^{\text{location}} \otimes Q_t^{\text{scale}} \otimes Q_t^{\text{shape}}. \tag{4}$$

This brings the total number of regimes to $H = 8$ $(= 2^3)$. It is important to note that each type of parameter (location, scale, and shape) can change over time according to Markov processes, of which the transition probabilities may be potentially influenced by the lagged conditioning variables. However, the times of changes for a specific parameter are stochastically independent of the times of changes for another one since there is no reason to believe that all parameters of the distribution change at the same time.

I use an univariate probit model to measure the evolution of the unobservable regime s_t^k as follows:

$$\Pr[s_t^k = 1] = \Pr[s_t^{*,k} \geq 0], \tag{5}$$

where $s_t^{*,k}$ is a latent variable defined as:

$$s_t^{*,k} = \gamma_0^k + \gamma_1^k s_{t-1}^k + \gamma_{z,1}^k (1 - s_{t-1}^k) z_t + \gamma_{z,2}^k s_{t-1}^k z_t + u_t^k, \tag{6}$$

where u_t^k has the following distribution:

$$p(u_t^k) = \text{normal}(u_t^k | 0, 1), \tag{7}$$

where $\text{normal}(x|\mu, \Sigma)$ denotes the normal distribution of x with mean μ and variance Σ. The parameters $\gamma_{z,1}^k$ and $\gamma_{z,2}^k$ determine how conditioning variables affect the transition probabilities of regimes. Thus, the transition probabilities $q_{1,1}^k(z_t)$ and $q_{2,2}^k(z_t)$ are given by:

$$\begin{aligned} q_{1,1}^k(z_t) &= \Pr(s_t^k = 0 | s_{t-1}^k = 0, z_t] \\ &= \Pr[s_t^{*,k} < 0 | s_{t-1}^k = 0, z_t] \\ &= \Pr[u_t^k < -\gamma_0^k - \gamma_{z,1}^k z_t] \\ &= \Phi(-\gamma_0^k - \gamma_{z,1}^k z_t). \end{aligned} \tag{8}$$

$$\begin{aligned} q_{2,2}^k(z_t) &= \Pr[s_t^k = 1 | s_{t-1}^k = 1, z_t] \\ &= \Pr[s_t^{*,k} \geq 0 | s_{t-1}^k = 1, z_t] \\ &= \Pr[u_t^k \geq -\gamma_0^k - \gamma_1^k - \gamma_{z,2}^k z_t] \\ &= 1 - \Phi(-\gamma_0^k - \gamma_1^k - \gamma_{z,2}^k z_t). \end{aligned} \tag{9}$$

with $\Phi(.)$ represents the cumulative distribution function for u_t^k. When $\gamma_{z,1}^k = 0$ and $\gamma_{z,2}^k = 0$, we have fixed transition probabilities, and the lagged conditioning variable becomes obsolete. The persistence of each regime is gauged by the transition probabilities. As a consequence, the

persistence of each business cycle regime depends on the previous business cycle regime, s_{t-1}^k, and on the lagged conditioning variable z_t. In other words, variations in z_t can potentially affect the expected duration of how long a regime will last.

For a skew-normal model with fixed parameters the likelihood can be easily evaluated and then combined with a prior distribution for the parameters. When dealing with a Markov-switching skew-normal model, it can be evaluated according to the Hamilton (1989)'s filter. Please refer to appendix A for a description of the evaluation of the likelihood. Since the posterior density function is very non-Gaussian, it is essential to find the posterior mode via an optimization routine. The estimate of the mode not only represents the most likely value (and thus the posterior estimate), but also serves as a crucial starting point for initializing different chains of MCMC draws.

The strategy to find the posterior mode is to generate a sufficient number of draws from the prior distribution of each parameter. Each set of points is then used as starting points to the `CSMINWEL` program, the optimization routine developed by Christopher A. Sims. Starting the optimization process at different values allows me to correctly cover the parameter space and avoid getting stuck in a "local" peak. Note, however, that I do not need to use a more complicated method for finding the mode like the blockwise optimization method developed by Sims, Waggoner, and Zha (2008). The authors employ a class of richly parameterized multivariate Markov-switching models in which the parameters are break into several subblocks, and then apply a standard hill-climbing quasi-Newton optimization routine to each block, while keeping the other subblocks constant, in order to maximize the posterior density. The size of the Markov-switching univariate model remains relatively small and allows me to employ a more standard technique.

Then, a Gibbs sampler procedure begins with setting parameters at the peak of the posterior density function. The MCMC sampling sequence involves a 6-block Gibbs sampler, in which I can generate in a flexible and straightforward manner alternatively draws from full conditional posterior distributions. The next section provides the details of the algorithm.

III.2. **Gibbs Sampler Procedure.** In the existing statistical literature, efficient posterior simulation algorithms have been applied to finite mixtures of skew-normal distributions. See, for example, Lin, Lee, Yen, and Chung (2007) and Frühwirth-Schnatter and Pyne (2010). My work differs from this literature along several dimensions. I assume that regime shifts evolve according to a Markov chain. Finite mixture models seems to be less suited for time series analysis as they consider unrealistically rapid switching regimes. By contrast, Markov-switching models can be seen as an extension of mixture models with a general solution to the problem of state persistence. Second, the MCMC algorithm developed in this paper is able to

directly generate draws of the shape parameters from a closed-form full conditional posterior distribution, and thus avoiding to employ a Random-walk Metropolis-hasting (RWMH) algorithm. Third, the method is able to deal with independently switches between the location, the scale and the shape parameters over time, while Frühwirth-Schnatter and Pyne (2010) allow only switches in a synchronized manner. Overall, the MCMC approach can be seen as an extension of Albert and Chib (1993).[4]

A MCMC simulation method is employed to approximate the joint posterior density $p(\theta, \Xi_T, S_T|Y_T)$, where $\theta = (\mu, \sigma, \alpha, \gamma)$, $S_t = [s_1, \ldots, s_t]$, and $\Xi_t = [\xi_1, \ldots, \xi_t]$ for $t \geq 1$. Define $\mu = (\mu_1, \mu_2)'$, $\sigma = (\sigma_1, \sigma_2)'$, $\alpha = (\alpha_1, \alpha_2)'$ and $\gamma = (\gamma^{\text{location}}, \gamma^{\text{scale}}, \gamma^{\text{shape}})$, where $\gamma^k = (\gamma_0^k, \gamma_1^k, \gamma_{z,1}^k, \gamma_{z,2}^k)'$. Here, a key to Bayesian estimation of a Markov-switching skew-normal model is to apply a stochastic representation of equation (1) as follows:

$$y_t = \mu_{s_t^{\text{location}}} + \delta_{s_t^{\text{shape}}} \xi_t + \sqrt{1 - \delta^2_{s_t^{\text{shape}}}} \nu_t, \qquad t = 1, \ldots, T \tag{10}$$

where $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$, ξ_t and ν_t are random variables at time t defined, respectively, as follows:

$$\text{truncated-normal}(\xi_t|0, \sigma^2_{s_t^{\text{scale}}})_{\xi_t > 0} \qquad \text{and} \qquad \text{normal}(\nu_t|0, \sigma^2_{s_t^{\text{scale}}}), \tag{11}$$

with truncated-normal$(x|\mu, \Sigma)_{x>0}$ denotes the truncated normal distribution of x with mean μ and variance Σ and truncated to positive values.

Because I consider a Bayesian approach to model (10) and (11), I now explicit the prior. The prior on the set of parameters θ is given by:

$$p(\mu_\upsilon) = \text{normal}(\mu_\upsilon|\bar{\mu}_1, \bar{\mu}_2), \tag{12}$$

$$p(1/\sigma_\upsilon^2) = \text{gamma}(1/\sigma_\upsilon^2|\bar{\sigma}_1, \bar{\sigma}_2), \tag{13}$$

$$p(\alpha_\upsilon) = \text{normal}(\alpha_\upsilon|\bar{\alpha}_1, \bar{\alpha}_2), \tag{14}$$

$$p(\gamma^k) = \text{normal}(\gamma^k|\bar{\gamma}_1^k, \bar{\gamma}_2^k), \tag{15}$$

where $\upsilon \in \{1, 2\}$, $\bar{\mu}_1$, $\bar{\mu}_2$, $\bar{\sigma}_1$, $\bar{\sigma}_2$, $\bar{\alpha}_1$, $\bar{\alpha}_2$, $\bar{\gamma}_1^k$, and $\bar{\gamma}_2^k$ are the hyperparameters, and gamma$(x|\bar{x}_1, \bar{x}_2)$ denotes the gamma distribution with hyperparameters $\bar{x}_1$ and $\bar{x}_2$. As can be seen, I directly specify informative priors for the shape parameter α_υ rather than for δ_υ, the transformed shape parameters. When specifying priors for δ_υ, instead of α_υ, there is no closed form for the posterior distribution of δ_υ, and one must impose a non-informative prior (i.e., uniform distribution on a bounded interval between -1.00 and 1.00), and use a RWMH algorithm.

The new representation in (10) and (11) leads me to exploit the idea of Gibbs-sampling. Let $\theta_{\neq x}$ contain the model's parameters, except for x, $S_t^* = [s_1^*, \ldots, s_t^*]$, and $Z_t = [z_1, \ldots, z_t]$.

[4] Albert and Chib (1993) develop a Gibbs sampling for univariate models subject to Markov mean and variance shifts.

The MCMC sampling scheme at the (n)st iteration, for $n = 1, \ldots, N_1 + N_2$, consists of sampling from the following conditional posterior distributions

(1) $p\left(S_T^{(n)}|Y_T, \theta^{(n-1)}\right)$;
(2) $p\left(S_T^{*(n)}, \gamma^{(n)}|S_T, Z_T\right)$;
(3) $p\left(\Xi_T^{(n)}|Y_T, S_T^{(n)}, \theta^{(n-1)}\right)$;
(4) $p\left(\mu_v^{(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\mu}^{(n-1)}\right)$;
(5) $p\left(1/\sigma_v^{2(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\sigma}^{(n)}\right)$;
(6) $p\left(\alpha_v^{(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\alpha}^{(n)}\right)$.

A few items deserve discussion. First, simulation from the conditional posterior density $p\left(S_T^{(n)}|Y_T, \theta^{(n-1)}\right)$, given Z_T and θ, is standard and in closed form. Second, simulation from the conditional posterior density $p\left(S_T^{*(n)}, \gamma^{(n)}|S_T^{(n)}, Z_T\right)$ is a two-step procedure. Simulation of $S_T^{*(n)}$ are based on equation (6) by generating u_t from an appropriate truncated standard normal distribution. Then, conditional on $S_T^{*(n)}, S_T^{(n)}, Z_T$, simulation of γ is available in closed form and follows a normal distribution using equation (6), which represents a simple linear regression model. Once values for $\gamma_0^{(n)}, \gamma_1^{(n)}, \gamma_{z,1}^{(n)}, \gamma_{z,2}^{(n)}$ are generated, one can easily compute the transition probabilities $q_{1,1}^k$ and $q_{2,2}^k$ using the normal cumulative density function in equations (8) and (9), respectively. Fourth, simulation from the conditional posterior density $p\left(\Xi_T^{(n)}|Y_T, S_T^{(n)}, \theta^{(n-1)}\right)$, given Y_t, S_t and θ, is available in closed form due to the stochastic representation of the Markov-switching model through normal and truncated-normal variables. Fifth, simulations from the conditional posterior densities $p\left(\mu_v^{(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\mu}^{(n-1)}\right)$ and $p\left(1/\sigma_v^{2(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\sigma}^{(n)}\right)$ reduces to Bayesian inference for Markov-switching models with known allocations, S_T. Finally, simulation from the conditional posterior density $p\left(\alpha_v^{(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\alpha}^{(n)}\right)$ is in closed form, and follows an unified skew-normal distribution introduced by Arellano-Valle and Azzalini (2006). Overall, the Gibbs sampling procedure is appealing as one can generate in a flexible and straightforward manner alternatively draws from full conditional posterior distributions, which are all in closed form. Appendix B provides the computational details for each conditional posterior distribution.

The sampler begins with setting parameters at the peak of the posterior density function. I collect N_1+N_2 draws of the MCMC sequence and keep only the last N_2 values. The only computational complication involves the simulation from the posterior distribution of α, which requires to sample from a truncated multivariate normal distribution. I use the minimax tilting method proposed by Botev (2017) for exact independently and identically distributed

data simulation from the truncated multivariate normal distribution.[5] The method is an excellent algorithm designed for extremely fast simulation.

IV. Empirical Results

In this section, I use the model presented in the previous section to assess whether economic and financial conditions provide timely warning signals of downside risk in economic activity. Toward this aim, I consider (quarter-on-quarter) GDP growth as variable of interest, y_t, and the lagged GDP growth and CISS index as conditioning variables, z_t. The overall sample period is the first quarter of 1999 to the fourth quarter of 2019. Section IV.1 presents the specification for inference. Section IV.2 reports the main results. Finally, Section IV.3 performs model comparison.

IV.1. **Specification for inference.** The priors are defined on the left-hand side of Table 1. A few of them deserve further discussion. Regarding the parameters governing the distribution, for μ_v I choose a normal prior with the mean 0.00 and the standard deviation 2.00. The prior for the scale parameter, $1/\sigma_v^2$, follows a gamma distribution, with hyperparameters $\bar{\sigma}_1$ and $\bar{\sigma}_2$ equal to one. This prior is thus very dispersed and covers a large parameter space. The prior for the shape parameter, α_v, has a normal density with the mean 0.00 and the standard deviation 4.00. It may be worth noting that I impose the exact same prior across regimes, so that the differences in parameters between regimes result more from data (i.e., the likelihood) rather than priors.

A little more explanation is required for the prior on the parameters determining transition probabilities. The priors for the intercepts of the transition probabilities, γ_0^k and γ_1^k, are allowed to differ across Markov processes. Their values are chosen to reflect a prior belief when conditioning variables assume values close to the sample mean ($z_t = 0$) or when the slope coefficient is zero. Regarding the process s^{location}, $\gamma_0^{\text{location}}$ and $\gamma_1^{\text{location}}$ are both given normal priors centered on -0.70 and 2.20, respectively, with tight standard deviations 0.10. This is consistent with a prior duration of 4 and 14 quarters for low and high GDP growth regimes, respectively. Regarding the process s^{scale}, the prior for γ_0^{scale} and γ_1^{scale} are centered on -0.60, and 2.30, respectively, with tight standard deviations 0.10. This implies a prior duration of 3.5 and 20 quarters, for high and low volatility GDP growth regimes, respectively. Regarding the process s^{shape}, the prior for γ_0^{shape} and γ_1^{shape} are centered on -1.00, and 2.00, respectively, with tight standard deviations 0.10, implying a prior duration for each regime of about 6 quarters. In the robustness section, I show that other values for standard deviation do not alter the posterior estimates.

[5]The Matlab function is available at https://fr.mathworks.com/matlabcentral/fileexchange/53792-truncated-multivariate-normal-generator.

TABLE 1. Prior and Posterior Distributions.

	Prior			Posterior					
Coefficient	Density	para(1)	para(2)	Mean	Median	[16;	84]	[5;	95]
Skew-normal parameters									
$\mu(s^{\text{location}}=1)$	N	0.00	2.00	0.0023	−0.0184	−0.1412	0.1631	−0.2244	0.2683
$\mu(s^{\text{location}}=2)$	N	0.00	2.00	0.5548	0.5591	0.5181	0.5947	0.4735	0.6237
$1/\sigma^2(s^{\text{scale}}=1)$	G	1.00	1.00	0.3975	0.3546	0.1694	0.6199	0.0915	0.8588
$1/\sigma^2(s^{\text{scale}}=2)$	G	1.00	1.00	8.3972	8.2792	6.7501	10.0198	5.8756	11.2963
$\alpha(s^{\text{shape}}=1)$	N	0.00	4.00	−4.5595	−4.2092	−6.6372	−2.5223	−8.5680	−1.7908
$\alpha(s^{\text{shape}}=2)$	N	0.00	4.00	3.9976	3.6019	1.7675	6.2632	0.8921	8.4470
Transition probability parameters, s^{location}									
$\gamma_0^{\text{location}}$	N	−0.70	0.10	−0.6818	−0.6811	−0.7778	−0.5849	−0.8390	−0.5247
$\gamma_1^{\text{location}}$	N	2.20	0.10	2.2182	2.2185	2.1234	2.3146	2.0559	2.3773
$\gamma_{\text{gdp},1}^{\text{location}}$	N	0.00	2.00	−0.2135	−0.2798	−1.1643	0.8072	−1.8767	1.6764
$\gamma_{\text{ciss},1}^{\text{location}}$	N	0.00	2.00	−0.9005	−0.8814	−2.8555	1.0178	−4.0274	2.2459
$\gamma_{\text{gdp},2}^{\text{location}}$	N	0.00	2.00	0.7729	0.7172	−0.7047	2.2596	−1.4481	3.3044
$\gamma_{\text{ciss},2}^{\text{location}}$	N	0.00	2.00	−1.6524	−1.6951	−3.4728	0.1557	−4.5842	1.4387
Transition probability parameters, s^{scale}									
γ_0^{scale}	N	−0.60	0.10	−0.5527	−0.5539	−0.6452	−0.4596	−0.7076	−0.3944
γ_1^{scale}	N	2.30	0.10	2.3405	2.3404	2.2452	2.4351	2.1866	2.4964
$\gamma_{\text{gdp},1}^{\text{scale}}$	N	0.00	2.00	0.6217	0.3599	−0.6669	1.9868	−1.1373	3.1425
$\gamma_{\text{ciss},1}^{\text{scale}}$	N	0.00	2.00	0.0928	0.1694	−1.9228	2.0685	−3.3316	3.3073
$\gamma_{\text{gdp},2}^{\text{scale}}$	N	0.00	2.00	0.7596	0.7725	−0.1492	1.6494	−0.7938	2.3090
$\gamma_{\text{ciss},2}^{\text{scale}}$	N	0.00	2.00	−0.9188	−0.9179	−2.5372	0.7158	−3.6515	1.6711
Transition probability parameters, s^{shape}									
γ_0^{shape}	N	−1.00	0.10	−1.0229	−1.0217	−1.1162	−0.9284	−1.1744	−0.8672
γ_1^{shape}	N	2.00	0.10	1.9918	1.9927	1.8932	2.0894	1.8316	2.1550
$\gamma_{\text{gdp},1}^{\text{shape}}$	N	0.00	2.00	1.9169	1.8196	0.6251	3.2171	−0.0014	4.1702
$\gamma_{\text{ciss},1}^{\text{shape}}$	N	0.00	2.00	0.1757	0.1821	−1.4747	1.8064	−2.5686	2.8250
$\gamma_{\text{gdp},2}^{\text{shape}}$	N	0.00	2.00	−0.0364	−0.1945	−1.0303	0.8760	−1.5683	2.1912
$\gamma_{\text{ciss},2}^{\text{shape}}$	N	0.00	2.00	−0.6103	−0.6446	−2.5147	1.2730	−3.7037	2.5457

Note: N stands for Normal, and G for Gamma distributions. The 16 percent and 84 percent demarcate the bounds of the 68 percent probability interval, while the 5 percent and 95 percent demarcate the bounds of the 90 percent probability interval. Para(1) and Para(2) correspond to the means and standard deviations for Normal Distributions, and to the hyperparameters for Gamma Distributions.

Finally, the slope coefficients, $\gamma^k_{z,1}$ and $\gamma^k_{z,2}$, are constrained to be equal across Markov processes. I impose a prior centered on zero and with a standard deviation 2.00. This is meant to reflect that the researcher is skeptical of the leading properties of conditioning variables but does not rule it out at all.

I estimate and simulate the model with Markov shifts using the maximization and MCMC procedures developed in the previous section. In particular, the results shown in this paper are based on $N_1 + N_2 = 55,000$ draws from the Gibbs-sampling procedure. I discard the first $N_1 = 5,000$ draws as burn-in, and keep every 10-th draw in order to achieve an approximately independent sample.

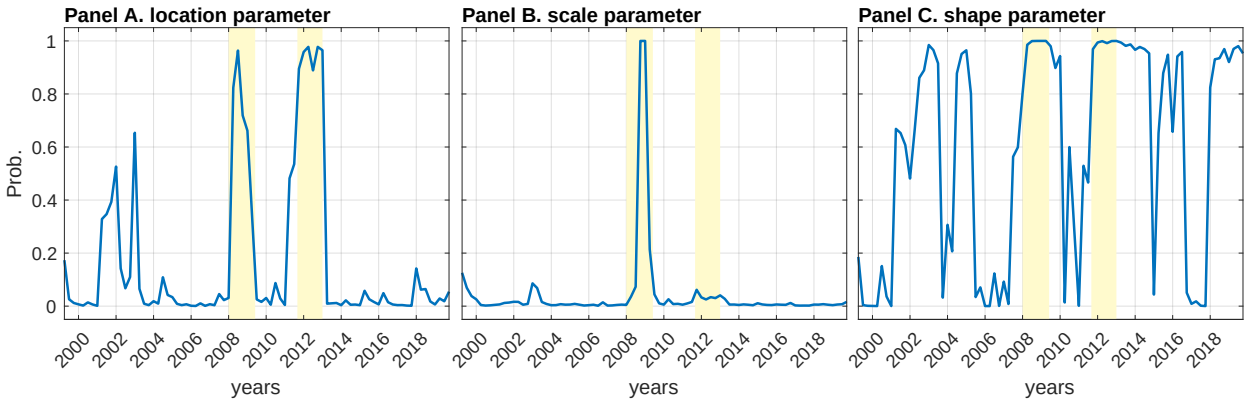
IV.2. **Model estimates.** On the right-hand side of Table 1, I report the posterior mean, and median with the 68 and 90 percent probability interval for each parameter of the estimated model. The first finding that is evident is the remarkable difference in the estimated parameters across regimes. For example, the median value of the location parameter in the first regime is about -0.02, while its value in the second regime is positive (0.56). Regarding the scale parameter, its value appears to be more than four times higher in the first regime. The shape parameter appears to be negative in the first regime, but positive in the second one (α_1=-4.21 and α_2=3.60 at the median). Another interesting aspect is that, for each parameter, the 68% and 90% probability intervals do not overlap between regimes. This is a signal that those parameters are well identified in the model and that a two-state specification is necessary to capture nonlinearities in the GDP growth.

The estimated elasticities of transition probabilities, $\gamma^k_{z,1}$ and $\gamma^k_{z,2}$, are a way to understand the relevance of changes in economic and financial conditions on the probabilities of switching regimes, and thus ultimately on the conditional distribution of GDP growth. Except for the case of $\gamma^{\text{shape}}_{gdp,2}$, the posterior median estimates of all elasticities are poorly estimated since their 90% probability intervals are very large and lie within both negative and positive regions. The posterior distribution of the elasticity $\gamma^{\text{shape}}_{gdp,2}$ has most of its mass on the positive side, suggesting that good economic conditions tend to anticipate positively skewed distributions. Regarding the posterior distribution of $\gamma^{\text{location}}_{ciss,2}$, although the posterior 90% probability interval does include zero, the corresponding 68% almost do not, and this could be interpreted as supporting the idea that financial conditions are empirically relevant to predict the GDP growth distribution, and more exactly its location. Looking at the posterior distributions of $\gamma^{\text{shape}}_{ciss,.}$, there is no evidence that financial conditions help to anticipate skewness, which is mainly driven by the time-varying shape parameter (see below). These estimates however do not provide a formal evaluation of both economic and financial conditions to predict GDP growth, and in particular, its tail risks. This is the objective of next sections.

Another interesting feature is that the parameters which describe transition parameters, γ_0^k and γ_1^k, are relatively better identified although the data are not able to steer much away the posterior medians of these parameters from their prior values. For example, for $\gamma_0^{\text{location}}$ and $\gamma_1^{\text{location}}$, we have -0.68 and 2.22 as posterior medians while the corresponding prior medians are 0.70 and 2.20. As a robustness check (see Section VII), I examine other degrees of tightness and show that they deliver similar estimates.

Figure 3 reports the (filtered) probabilities — evaluated at the mode — of being in the first regime at each period and for each Markov process.[6] One can see from the figure that the euro area economy has been characterized by numerous switches between regimes over time. Looking at the process governing the location parameter (Panel A), Regime 1 coincides remarkably well with the recessions dates declared by the CEPR's Business Cycle Dating Committee. This result is definitively in line with Hamilton (1989). Regarding the

FIGURE 3. Regime probabilities.



Note: Sample period: 1999.Q2 — 2019.Q4. Probabilities of being in Regime 1 produced from the baseline model. The location (Panel A), scale (Panel B) and shape (Panel C) parameters of the skew-normal distribution are allowed to change according to two-state independent Markov-switching processes. The yellow areas denote the CEPR recessions.

process governing the scale parameter (Panel B), one can see from the figure that the low-scale regime has been predominantly prevailed throughout the sample, with the exception of the Great Recession period. While nearing 0 during the early 2000s, the probability of the high-scale regime rapidly rises in the beginning of 2008. This regime covers largely the Great Recession of 2008-2009, but then it has never been in place again. This finding corroborates with Lhuissier (2017, 2018), who estimates richly parameterized multivariate

[6]Results are similar when reporting the smoothed probabilities in the sense of Kim (1994); i.e., full sample information is used in getting the regime probabilities at each date.

TABLE 2. Estimated transition matrices (when $z_t = 0$)

Location			Scale			Shape		
	$s_t^{\text{lo}} = 1$	$s_t^{\text{lo}} = 2$		$s_t^{\text{sc}} = 1$	$s_t^{\text{sc}} = 2$		$s_t^{\text{sh}} = 1$	$s_t^{\text{sh}} = 2$
$s_t^{\text{lo}} = 1$	0.7521 [0.7207;0.7817]	0.0621 [0.0476;0.0801]	$s_t^{\text{sc}} = 1$	0.7102 [0.6772;0.7407]	0.0370 [0.0277;0.0484]	$s_t^{\text{sh}} = 1$	0.8465 [0.8234;0.8679]	0.1667 [0.1351;0.2017]
$s_t^{\text{lo}} = 2$	0.2479 [0.2183;0.2793]	0.9379 [0.9199;0.9524]	$s_t^{\text{sc}} = 2$	0.2898 [0.2593;0.3228]	0.9630 [0.9516;0.9723]	$s_t^{\text{sh}} = 2$	0.1535 [0.1321;0.1766]	0.8333 [0.7983;0.8649]

Note: Posterior medians and [16th , 84th] percentiles are reported.

models for the euro area economy with regime changes in shock variances over time and associates the Great Recession to a high-volatility regime. Finally, regarding the process governing the shape parameter, Panel C suggests that the euro area economy has been characterized by numerous switches between the negative- and positive-shape regimes over time. In particular, the negative-shape regime has been sporadically prevailed in the early 2000s, and predominantly in 2008-2009, 2012-2015, and 2018-2019. Interestingly, although the shape parameter is not the only source of potential asymmetry in the distribution due to mixing probabilities, it appears that downside risks dominate a large part of the sample. The analysis of the next section will corroborate this finding.

Overall, the results show clearly that location, scale, and shape parameters do not switch at the same time, strengthening the specification of independent Markov-switching processes made in the baseline model. I perform a more formal check in the robustness section.

Table 2 reports the posterior transition probabilities, Q^k, at the sample mean for conditioning variables ($z_t = 0$). Regarding the posterior probabilities of transition matrix (Q^{location}) that governs time variation in the location parameter, it is apparent that the persistence of staying in each state is relatively high. The 90% probability intervals for q_{11}^{location} are 0.72 and 0.78, and those for q_{22}^{location} are 0.92 and 0.95, indicating that the first regime is much less persistent (an average duration of 4 quarters at the median) than the second regime (an average duration of 14 quarters at the median). Posterior means and medians are concentrated in tight ranges, reinforcing the estimated parameters. The pattern is similar for the transition matrix Q^{scale}, where the posterior duration of the first regime is shortened with respect to the second one. This is not very surprising since the regime prevailed only during the Great Recession period, as shown in Panel B of Figure 3. Finally, the transition matrix Q^{shape} reveals an average duration of about 6 quarters for each regime, implying recurrent downside and upside risks in the euro area economy.

IV.3. **Information Critera for alternative specifications.** 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

TABLE 3. Information criteria

	WAIC	standard errors
MS skew-normal model with GDP and CISS	34.2688	7.2745
MS skew-normal model with GDP	35.0700	7.4293
MS skew-normal model	33.6708	7.6450
MS Gaussian model	44.0676	7.6903

Note: Watanabe-Akaike information criteria (WAIC). WAIC $= -\log(\overline{lpd}) + \overline{p}$, where $\log(\overline{lpd})$ is the log pointwise predictive density, i.e., $\sum_{t=1}^{T} \log\left(\frac{1}{S}\sum_{s=1}^{S} p(y_i|\theta^s)\right)$ with S is the number of MCMC iteration, and $\overline{p}$ is the estimated effective number of parameters, i.e., $\sum_{t=1}^{T} V_{s=1}^{S}(\log(p(y_i|\theta^s))$, with V represent the sample variance. The standard error $se(\overline{elpd}) = \sqrt{(T * V_{t=1}^{T}\overline{elpd}, t)}$, where $\overline{elpd}, t = \log\left(\frac{1}{S}\sum_{s=1}^{S} p(y_i|\theta^s)\right) - \left(V_{s=1}^{S}\log p(y_i|\theta^s)\right)$.

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.

To compare the information content of the baseline model ("MS[7] skew-normal model with GDP and CISS"), I consider alternative Markov-switching specifications: MS skew-normal model with time-varying transition probabilities conditional only on GDP ("MS skew-normal model with GDP"); MS skew-normal with fixed time-varying probabilities ("MS skew-normal model"); and a standard Markov-switching Gaussian model with fixed time-varying probabilities ("MS Gaussian model").[8] Table 3 reports the value and the standard deviation of WAIC for each model. There are a number of best-fit models, all of them are MS skew-normal models, which delivers very similar fit. The WAICs for these models are higher by at least 10 on a log scale than that for the MS Gaussian model. Clearly, the inclusion of time-varying asymmetry in GDP growth is supported by the data. However, MS skew-normal models that incorporate additional information from economic and financial sectors does not help to improve the fit of the model.

[7] "MS" stands for Markov-switching.

[8] The MS Gaussian model is characterized by two independent Markov-switching processes; one governing the location parameter and one governing the scale parameter. Although not reported, the times of location and scale changes are very similar to those produced from MS skew-normal models.

V. Time variation in GDP growth distribution

In this section I explore the time-series evolution in the higher order moments, tail risks and predictive densities implied by the baseline model that incorporates both economic and financial variables. For comparison purposes, I also consider alternative Markov-switching models. All exercises are based on in-sample experiments.[9] Section V.1 studies time variation in the conditional first fourth moments. Section V.2 focuses on tail risks, and Section V.3 reports predictive densities to assess whether financial conditions provide a timely warning in the early stages of the 2008–09 recession.

V.1. **Time variation in euro area moments.** The previous section demonstrated that all parameters of the skew-normal distribution are time-varying, thus indicating that the GDP growth of the euro area is clearly not normally distributed. However, the analysis does not allow to directly inspect 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). For completeness, this section studies the evolution of the mean, variance, skewness, and kurtosis, with a particular emphasize on the skewness moment in order to provide an accurate characterization of the downside risk to GDP growth in the euro area.[10] Deriving the first four centered, conditional moments of the Markov-switching model is not straightforward. Timmermann (2000) characterizes the moments of the basic Markov switching model for the cases where the error term follows a t-distribution or a normal distribution. I extend the approach for the case where the error term is a skew-normal distribution. Appendix C provides the computational details of the first four moments.

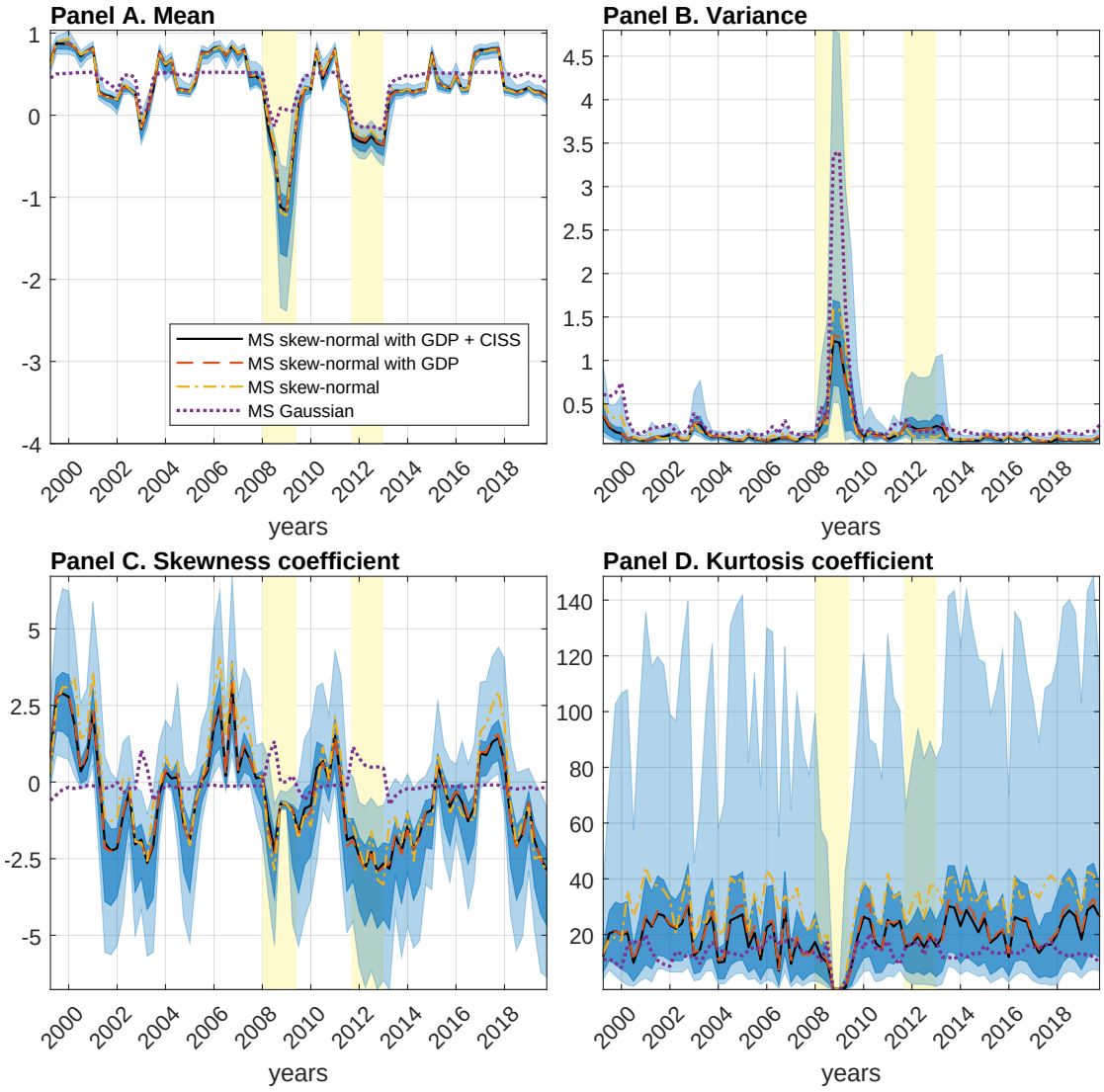
Figure 4 presents the time-varying moments of the distribution of GDP growth, drawn from the models under consideration. For visibility purposes, I report the median and probability bands for the baseline model that incorporates both financial and economic conditions, while I only report medians for the remaining models. As can be seen from all models, there is a clear cyclical pattern in the mean which declines toward the end of expansions and rise during the end of the recession periods. Note however that the simple MS Gaussian model is not able to produce larger fluctuations during the Great Recession with respect to other models. Turning to the second moment, GDP growth generates clear counterfactual patterns, with a

[9] I consider in-sample experiments. I do not need to resort to simulation, since the one-step-ahead Markov-switching skew-normal density is available in closed form.

[10] This moment is increasingly used to characterize risk. A growing number of theoretical and empirical studies have emphasized the importance of non-Gaussian shocks in macroeconomic models. Notable examples include Rietz (1988), Barro (2009), Barro and Ursúa (2012), Gabaix (2012), and Gourio (2012). They have suggested that rare disasters — rising from an asymmetric distribution of shocks — are a key driver of business cycle fluctuations, such as the Great Recession.

peak occurring during the Great Recession. Note however that when I introduce the shape parameter in the Markov-switching specification, variance appears to be much less important.

FIGURE 4. Moments.



Note: Sample period: 1999.Q2 — 2019.Q4. Historical path of first four moments — mean, variance, skewness, and kurtosis — of euro area GDP growth for four models (at the median): MS skew-normal with GDP + CISS (baseline model), MS skew-normal with GDP, MS skew-normal, MS Gaussian. The blue areas denote the 68% and 90% error bands for the baseline model. The yellow areas denote the CEPR recessions.

Regarding higher order moments, the coefficient of skewness follows a pronounced cyclical path with negative conditional skewness in the early recession stage. The source of asymmetry is not only due to time variation in the shape parameter, but also from the mixture feature

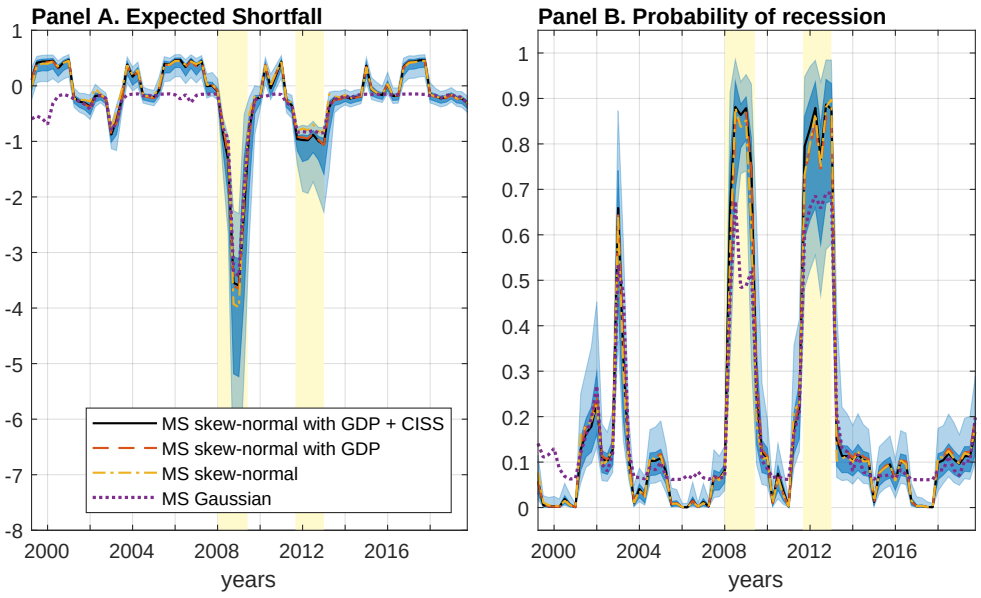
of the model. More broadly, negative skewness matches well the frequent occurrence of low values of GDP growth. The difference between the second and third moment is also clear, meaning that volatility is an insufficient measure of risk. The sample value of the skewness coefficient is -0.22, which results from large swings between positive and negative skews for all time periods. Interestingly, the skewness tends to decrease in anticipations of recessions, a feature that is also observable for the U.S. economy (e.g., De Polis, Delle Monache, and Petrella, 2020). To sum up, the skewness is procyclical and the time-varying business cycle asymmetry is very much in evidence. Interestingly, the simple MS Gaussian model is not able to produce reasonnable and realistic pattern in skewness, which tends to increase prior and during recession phases. Finally, the coefficient of excess kurtosis has a tendancy to rapidly decline prior and during the Great Recession, meaning that the distribution was relatively flat. For the rest of the sample, it remains relatively stable, with short-lived oscillations.

Regarding the posterior uncertainty around moments, the mean appears to be much more precisely estimated than the variance, skewness, and kurtosis. The historical path of kurtosis is even more imprecisely estimated. As can be seen from the figure, the implied uncertainty about the fourth moment is extremely large, with values ranging between 2 and 140 for almost all time periods. Interestingly, both models that do no incorporate financial conditions produce similar patterns and the median moments lie within the bands of the baseline model, meaning that there is no informational advantage in providing financial variables within the model.

V.2. **Time variation in tail risk.** I now investigate inferences about the left tail of the growth distribution. As metrics for the downside risks I use the expected shortfall, which can be readily obtained within my framework. Let ES^{α}_{t+h} be the expected growth level for $y_{t+h} < VaR^{\alpha}_{t+h}$, with VaR denotes the Value at Risk, and corresponding to the $(100\alpha)^{\text{th}}$ percentile of the h-step ahead predictive distribution. Specifically, $ES^{\alpha}_{t+h} = \alpha^{-1} \int_0^{\alpha} VaR^{a}_{t+h|t} da$ The expected shortfall equals expected growth conditional on growth falling below the fifth percentile of its conditional distribution. It therefore characterizes the severity of a recession should it materialize.

Panel A of Figure 5 shows the 5% expected shortfall for the four estimated models. Clearly, all models report similar patterns; they capture partially the build-up of downside risks ahead recessions. For example, the prediction of the ES falls to about -0.50% in the first quarter of the Great Recession (2008.Q1), and then decreases to reach a minimum of 4% in the early 2009. We see that ES varies substantially over time and is reasonably precisely estimated. The inclusion of financial-specific information in time-varying transition probabilities does not alter distributional forecasts, beyond the information contained in the real indicator.

FIGURE 5. Expected Shortfall and Probability of Recession.



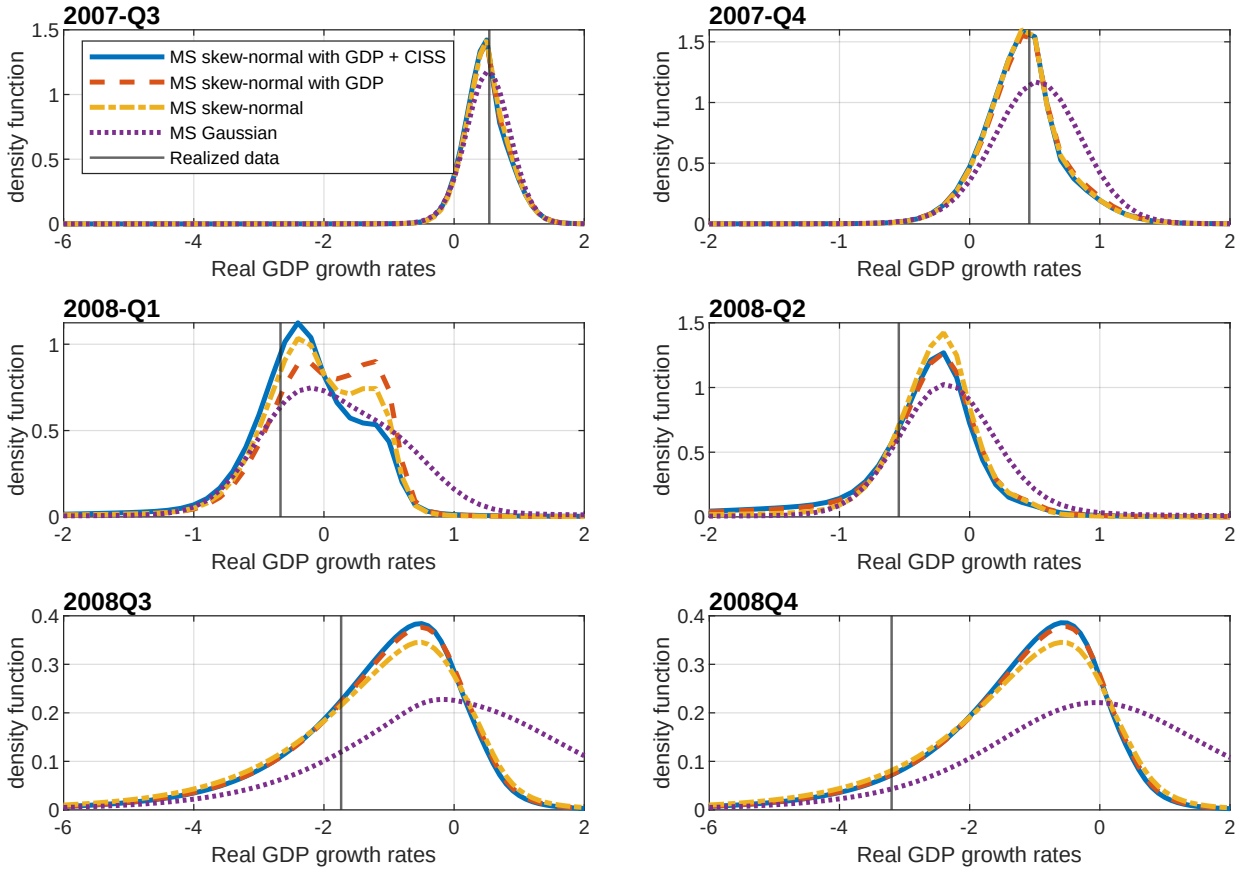
Note: Sample period: 1999.Q2 — 2019.Q4. Evolution of the 5% expected shortfall and the probability of recession for four models (at the median): MS skew-normal with GDP + CISS (baseline model), MS skew-normal with GDP, MS skew-normal, MS Gaussian. The blue areas denote the 68% and 90% error bands for the baseline model. The yellow areas denote the CEPR recessions.

Inference about the left tail of the growth distribution can also be investigated through the probability of a recession (that is, negative growth in the following quarter), as shown in Panel B of Figure 5. The recession probability varies substantially over time with great uncertainty. The MS skew-normal models produce a similar assessment of recession risk, with peaks matching the CEPR recession dates. The probability increase ahead of recessions, though in a modestly way, and not sufficiently to warn against an imminent economic downturn. At the late recession stage, the probability sharply declines and is below 10% just a quarter after the end of recessions as dated by the CEPR. Finally, the assessment of recession risk appears to be less precisely estimated for the MS Gaussian model.

V.3. **Predictive densities around the Great Recession.** The Great Recession hit the world economy unexpectedly. Sharp recessions in many parts of the world was triggered by capital losses after the Lehman Brothers event. It has been very difficult for traditional forecasting models to produce accurate forecasts for the evolution of GDP growth. This provides an interesting study case for the analysis of this paper. In this section, I ask whether

the model that includes financial conditions would have produced more accurate in-sample forecasts.

FIGURE 6. Predictive Densities.



Note: Predictive densities in the period of the Great Recession for four models (at the median): MS skew-normal with GDP + CISS (baseline model), MS skew-normal with GDP, MS skew-normal, MS Gaussian.

To provide a full characterization of risks of GDP growth around the Great Recession, I construct means of the predictive densities of GDP growth. Since I consider one-step ahead forecasts, I do not need to resort to simulation because the one-step-ahead Markov-switching skew-normal density is available in closed form.[11] Insights into how much the conditional

[11]More specifically, this means computing

$$p(y_t|Y_{t-1},\theta) = \sum_{i=1}^{H} \Pr[s_t = i] \frac{2}{\sigma_i} \phi\left(\frac{y_t - \mu_i}{\sigma_i}\right) \Phi\left(\alpha_i \frac{y_t - \mu_i}{\sigma_i}\right). \tag{16}$$

density of GDP growth varies quarter to quarter can be gained in Figure 6, which plots the sequence of quarterly densities during the Great Recession period 2007.Q3 to 2008.Q4.

MS skew normal models appear to better predict outcomes than MS Gaussian model, as they assign higher probability mass to economic downturn. At the end of the crisis, these models generate lowers conditional mean, higher conditional volatility and lower (negative) conditional skewness, which tends to attach higher likelihood to economic downturn. Nevertheless, all models seem to do poorly in capturing the shift in GDP growth for the last quarters of 2008. The results also reveal that the information content of financial conditions does not provide valuable additional information to predict GDP growth during the Great Recession. This is clearly in line with Plagborg-Møller, Reichlin, Ricco, and Hasenzagl (2020) which report that financial variables seems to have little informational advantage to predict the Great Recession in the U.S. economy.

In the next section, through a variety of forecast metrics, I provide a more systematic evaluation of the distributional forecast accuracy by performing out-of-sample forecasting exercises.

VI. Out-of-Sample Forecasting

In this section, I evaluate the out-of-sample performances of MS skew-normal models using a variety of forecast metrics, for one-quarter ahead prediction. The main objective is to assess whether conditioning on financial predictors leads to forecast improvements. Since I am studying macroeconomic downside risks, particular attention is paid on tail risk metrics.

I conduct out-of-sample backtesting exercises by using the proposed methodology in real time, with the caveat that I use final revised data only. The out-of-sample forcasting procedure is straightforward. I produce predictive distributions recursively, starting with the estimation sample that ranges from the 1999.Q2 to 2010.Q2. More precisely, using data from 1999.Q2 to 2010.Q2, I estimate the predictive distribution for 2010.Q3 (one quarter ahead). I then iterate the same procedure, expanding the estimation sample, one quarter at a time, until the end of the sample (2019.Q4). At each iteration, I repeat the estimation steps of Section III, maximizing the posterior distribution, running the Gibbs sampler, and computing forecast metrics. The online Appendix D presents the technical details.

I consider three types of metrics: point, density and tail risk. I assess the point forecast, defined as the median of the predictive distributions, via the traditional mean square forecast error (MSFE). Density forecasts accuracy is evaluated via the predictive log-score. To compute the score with models, 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).[12] Regarding the accuracy of the tail risk forecasts, I consider three measures of the lower tail quantile estimate at the 5 percent quantile along the lines of Carriero, Clark, and Marcellino (2020) and De Polis, Delle Monache, and Petrella (2020). The quantile is estimated as the associated percentile of the simulated predictive distribution. The first accuracy measure is a simple coverage measure for the interval forecast: the frequency with which real-time outcomes for GDP growth fall below the 5 percent quantile of the forecast distribution. The second is the quantile score, also known as the tick loss function (e.g., Giacomini and Komunjer, 2005).[13] The third is the joint value at risk-expected shortfall (VaR-ES) score from Fissler, Ziegel, and Gneiting (2015).[14]

Table 4 reports the predictive performance of models under consideration. To facilitate comparisons, except in the case of the coverage rates, the results are reported as relative to those for the MS Gaussian model. In particular, I report MSFE as a ratio of the MS Gaussian model's MSFE, and the quantile score as a ratio of the corresponding MS Gaussian model's score. A ratio below one means the model of interest is more accurate. I report the log predictive and VaR-ES scores as differences with respect to the MS Gaussian model, so that a differential above zero means the model of interest is more accurate. To gauge statistical significance, I estimate Diebold and Mariano (1995)-West (1996) t-tests for equality of the average loss (with loss defined here as squared error, log-score, quantile score, VaR-ES score). I also compute t-tests for the empirical coverage rate equaling the nominal rate of 5%. In the table, differences in accuracy that are statistically different from zero are denoted by "*", corresponding to significance levels of 5%. Significance of the test follows from Giacomini and White (2006)'s critical values.

In terms of point forecasting and density forecasts, all MS skew-normal models are broadly similar in accuracy. For point forecast, they generate lower MSFE values than the MS

[12]The Matlab function is available at https://fr.mathworks.com/matlabcentral/fileexchange/14034-kernel-density-estimator.

[13]The quantile score is computed as

$$QS_{\alpha,t+h} = (y_{t+h} - Q_{\alpha,t+h})\left(\alpha - \mathbb{1}_{(y_{t+h} \leq Q_{\alpha,t+h})}\right),$$

where y_{t+h} is the actual outcome for GDP growth, $Q_{\alpha,t+h}$ is the forecast quantile at quantile $\alpha = 0.05$, and the indicator function $\mathbb{1}_{(y_{t+h} \leq Q_{\alpha,t+h})}$ has a value of 1 if the outcome is at or below the forecast quantile and 0 otherwise.

[14]The joint VAR-ES score is computed as

$$\begin{aligned} S_{\alpha,t+h} =& Q_{\alpha,t+h}\left(\mathbb{1}_{(y_{t+h} \leq Q_{\alpha,t+h})} - \alpha\right) - y_{t+h}\mathbb{1}_{(y_{t+h} \leq Q_{\alpha,t+h})} \\ &+ \frac{e^{\mathrm{ES}_{\alpha,t+h}}}{1 + e^{\mathrm{ES}_{\alpha,t+h}}}\left[\mathrm{ES}_{\alpha,t+h} - Q_{\alpha,t+h} + \alpha^{-1}\left(Q_{\alpha,t+h} - y_{t+h}\mathbb{1}_{(y_{t+h} \leq Q_{\alpha,t+h})}\right)\right] + \ln\frac{2}{1 + e^{\mathrm{ES}_{\alpha,t+h}}}, \end{aligned}$$

where $\mathrm{ES}_{\alpha,t+h}$ denotes the expected shortfall forecast at quantile α.

TABLE 4. Accuracy of out-of-sample forecasts of GDP growth

Model	(a) MSFE	(b) Log-Score	(c) Tail risk scores		
			5% coverage	Quantile	VaR-ES
MS skew-normal with GDP and CISS	0.7621 (0.0973)	0.0755 (0.3449)	0.0295 (0.3101)	1.2448 (0.1179)	0.0414 (0.3929)
MS skew-normal with GDP	0.7691 (0.1048)	0.0814 (0.2916)	0.0274 (0.2629)	1.2662 (0.0930)	0.0541 (0.2652)
MS skew-normal	0.8887 (0.3356)	0.0943 (0.1943)	0.0214 (0.0506)	0.9803 (0.8393)	0.0076 (0.8026)

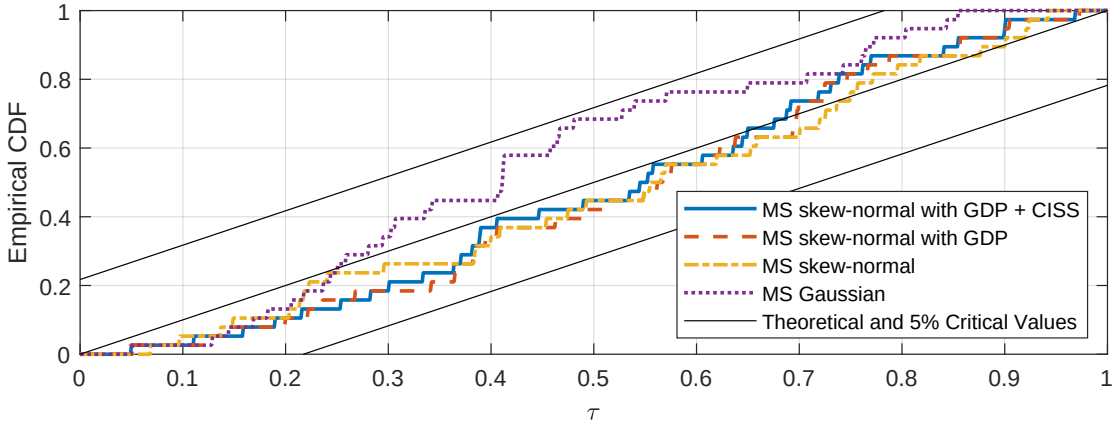
Note: The table reports average forecast metrics relative to the MS Gaussian model, except in the case of the 5 percent coverage rates. I use ratios for the MSFE, Quantile, and differences for Log-Score and VaR-ES. Ratios smaller than 1, and positive values of the log-score differences indicate that the model performs better than MS Gaussian benchmark. Values in parentheses report the p-values of the Diebold and Mariano (1995)-West (1996) test statistic for equal predictive accuracy. "*" denotes significance at the 5% level. Significance of the test follows from Giacomini and White (2006)'s critical values.

Gaussian model although none of the test statistics is significant at 5% level. For density forecast, all models tend to predict better the overall density. Clearly, there is substantial gain in the forecasting accuracy of the models featuring time-varying asymmetric distributions. However, for both metrics, the inclusion of financial-specific information does not improve the forecasting performance of the model.

Before turning to the tail risk scores, I provide an analysis of the calibration of the predictive distribution by looking at the properties of the probability integral transforms (PITs), which measures the percentage of observations that are below any given quantile α. Results are presented in Figure 7. Following Rossi and Sekhposyan (2019), I report confidence bands around the 45-degree line to account for sample uncertainty. For all models, the empirical distribution of the PITs lies within the confidence bands for any given quantile, suggesting that all models generate robust predictive distributions. Furthermore, PITs for the model conditional on both economic and financial conditions do not do better than others.

I then turn to the main focus of the paper, namely the assessment of tail risk predictions. To do so, I look at the following metrics: 5% coverage rates, quantile score, and VaR-ES score results, as show in the last column of Table 4. As to the 5% coverage measure, all models produce coverage rates reasonably close to the nominal rates, and none of the departures from 5 percent coverage appear to be statistically significant. Regarding quantile and VaR-ES scores, they deliver opposite results. The introduction of Markov shifts in the shape parameter outperforms (underperforms) the forecast accuracy of the model with respect

FIGURE 7. Probability Integral Transforms (PITs).



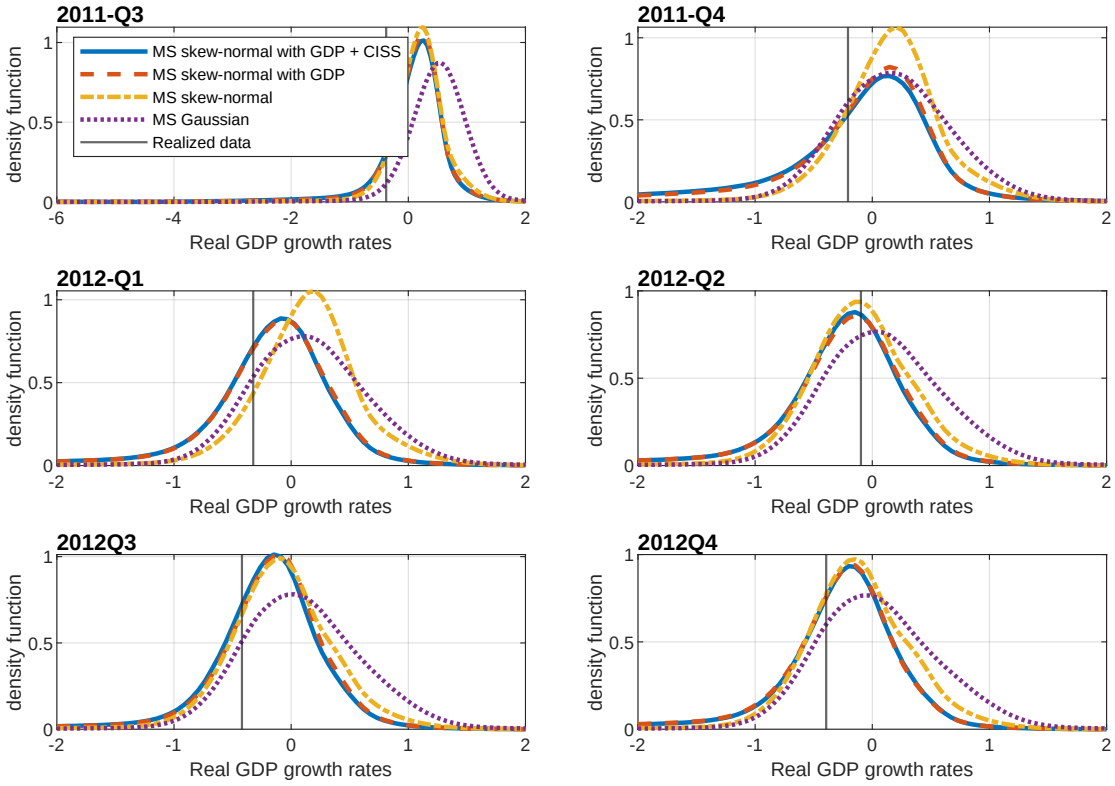
Note: Calibration of the predictive distribution by means of the probability integral transforms (PITs). The closer the empirical cumulative distribution of the PITs is to the 45 degrees line, the better the model is calibrated. Critical values are obtained as in Rossi and Sekhposyan (2019) to acccount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

to the benchmark MS Gaussian model in the case of the VaR-ES (Quantile) score. Note however that none of them are economically significant. Finally, looking at the three metrics, there is no significant gain in tail risk predictions in the model that include financial-specific information.

As an illustration, I now zoom in on the sovereign debt crisis, placing particular focus on the ability of models to anticipate the build up in downside risk ahead of the crisis. Figure 8 shows the sequence of quarterly predictive densities of GDP growth around this period. All models seem to do poorly in capturing the shift in economic conditions. This is even worst when considering the MS Gaussian model. While the MS Gaussian model performs particularly pooly, the three MS skew-normal models yield very similar predictive distributions.

I have run a similar exercise with the Covid-19 crisis, as reported in the online Appendix F. Using the same real-time estimation approach, I compute here the predictive distribution of GDP for the first quarter of 2020. Clearly, financial variables do not provide useful timely information about the COVID-19 downturn. A caveat of the analysis is the use of quarterly data, while financial variables are often available at higher frequency. Considering a real-time nowcasting exercise with high-frequency data might be more accurate to predict the COVID-19 downturn, as shown in Ferrara, Mogliani, and Sahuc (forthcoming).

FIGURE 8. Predictive densities around the sovereign debt crisis.



Note: One-period ahead predictive densities around the sovereign debt crisis for four models (at the median): MS skew-normal with GDP + CISS (baseline model), MS skew-normal with GDP, MS skew-normal, MS Gaussian.

VII. Robustness

I study a number of other relevant specifications in order to assess the robustness of results. Section VII.1 examines how the main results change when I modify the tightness of the prior. Section VII.2 studies models in which location, scale, and shape parameters switch at the same time. Finally, Section VII.3 compares in-sample forecasts to out-of-sample forecasts. All of these exercises reinforce the findings in the previous sections. These results of this section are available in the online appendix E.

VII.1. **Prior sensitivity.** In the baseline specification, I showed that the posterior means of the transition probability parameters, γ_0 and γ_1, are very close to their prior means. This may be not very surprising given that I impose relatively tight normal priors (i.e., standard deviations are equal to 0.10). In this section, other degrees of tightness for the priors were examined to determine if they deliver different outcomes. Specifically, I examined two level of standard deviations: 0.25 and 0.50. As shown in the online Appendix E.1, changes in the

prior tightness do not affect posterior distributions; prior means and posterior means remain relatively close. Furthermore, the evolution of conditional first fourth moments and tail risks produced with these alternative prior tightness remains remarkably similar to those produced from the baseline specification.

VII.2. **Synchronized vs independent chains.** The results so far assume that the times of changes for a specific parameter are stochastically independent of the times of changes for another one. In this section, 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 synchronized-chain models with that of independent-chains models.

In the online Appendix E.2, I display the value and the standard deviation of WAIC for each independent-chains model relative to their synchronized-chains counterparts. Clearly, all independent-chains models outperform synchronized-chains models since the estimated differences in their expected predictive accuracy are all negative. This is more clearly appreciated when taking into account the uncertainty (in terms of standard errors) of WAIC.

Overall, I conclude that independent-chains models are a better description of GDP growth than synchronized-chains models. These results are in line with Sims (2001) and Lhuissier and Zabelina (2015) showing that that making the transitions of variance and coefficient regimes independent delivers the best fit.

VII.3. **In-sample vs out-of-sample forecasts.** I study the robustness of the results shown so far by comparing the in-sample results with their out-of-sample counterparts. Results for this exercise are presented in the online Appendix E.3. In particular, I report selected percentiles computed from the baseline model that incorporates both economic and financial information using the full sample (in-sample) and recursively (out-of-sample). I show that the in-sample and out-of-sample estimates of the conditional distribution of future GDP growth are very remarkably similar, except for the period around the sovereign debt crisis. Indeed, the lower percentiles of GDP growth produced from out-of-sample estimates tend to decline much than those produced from in-sample estimates.

Furthermore, an evaluation of the distributional forecast accuracy (point, density and tail risk) in the in-sample experiments reveals that once again, the model incorporating the financial variable seems to have no advantage over other models under consideration, thus confirming the out-of-sample results.

VIII. Conclusion

The main goal in this paper was to examine the potential role of financial conditions in estimating the moments of the conditional distribution of GDP growth of the euro area and in predicting tail risks. To do so, I developed a regime-switching skew-normal model with time-varying transition probabilities. I relied on Bayesian methods to estimate the model and developed a Gibbs sampler which leads to alternatively draws from full conditional posterior distributions.

Model estimates provided evidence in support of a procyclical skewness in the GDP growth of the euro area. The skewness tends to decline ahead and during recessions, a feature that is also observable for the U.S. economy. However, the inclusion of financial-specific information in time-varying probabilities helps poorly to predict different features of the GDP growth distribution, including conditional skewness. Furthermore, through out-of-sample forecasting exercises, I showed that financial conditions cannot be seen as a warning signal of downside risks in GDP growth.

Extending Markov-switching skew-normal models to a multivariate framework would seem to be a natural next step. Another area of future work would be to relax the assumption of exogeneity of regime switching (i.e., the development of economic and financial conditions are exogenous to the model) in order to better understand the sources of changes in the conditional distribution of GDP growth. As such, the works by Kim, Piger, and Startz (2008) and Chang, Choi, and Park (2017) on endogenous Markov-switching univariate models could then be used in this direction. All in all, I believe those extensions certainly represent an interesting avenue for future research and would be suited to a variety of economic problems.

References

Adrian, T., and N. Boyarchenko (2012): "Intermediary leverage cycles and financial stability," Discussion paper.

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

Adrian, T., F. Grinberg, N. Liang, and S. Malik (2018): "The Term Structure of Growth-at-Risk," Hutchins center working paper n.42.

Albert, J. H., and S. Chib (1993): "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts," *Journal of Business & Economic Statistics*, 11(1), 1–15.

Arellano-Valle, R. B., and A. Azzalini (2006): "On the Unification of Families of Skew-Normal Distributions," *Scandinavian Journal of Statistics*, 33(3), 561–574.

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.

Bernanke, B., and M. Gertler (1989): "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review*, 79(1), 14–31.

Bernanke, B., M. Gertler, and S. Gilchrist (1999): "The Financial Accelerator in a Quantitative Business Cycle Framework," *Handbook of Macroeconomics*, pp. 1341–1393.

Botev, Z. I. (2017): "The normal law under linear restrictions: simulation and estimation via minimax tilting," *Journal of the Royal Statistical Society Series B*, 79(1), 125–148.

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

Brunnermeier, M., D. Palia, K. Sastry, and C. A. Sims (forthcoming): "Feedbacks: Financial Markets and Economic Activity," *American Economic Review*.

Brunnermeier, M. K., and Y. Sannikov (2014): "A Macroeconomic Model with a Financial Sector," *American Economic Review*, 104(2), 379–421.

Caldara, D., D. Cascaldi-Garcia, P. Cuba-Borda, and F. Loria (2020): "Understanding Growth-at-Risk: A Markov Switching Approach," Presented at the philaldephia fed conference on real-time data analysis, methods and applications.

Canale, A., E. Pagui, and B. Scarpa (2016): "Bayesian Modeling of University First-year Students' Grades after Placement Test," *Journal of Applied Statistics*, 43(16), 3015–3029.

Carriero, A., T. E. Clark, and M. Marcellino (2020): "Capturing Macroeconomic Tail Risks with Bayesian Vector Autoregressions," Working Papers 202002R, Federal Reserve Bank of Cleveland.

Carter, C. K., and R. Kohn (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81(3), 541–553.

Chang, Y., Y. Choi, and J. Y. Park (2017): "A new approach to model regime switching," *Journal of Econometrics*, 196(1), 127 – 143.

Christoffersen, P., S. Heston, and K. Jacobs (2006): "Option valuation with conditional skewness," *Journal of Econometrics*, 131(1-2), 253–284.

De Polis, A., D. Delle Monache, and I. Petrella (2020): "Modeling and Forecasting Macroeconomic Downside Risk," CEPR Discussion Papers 15109, C.E.P.R. Discussion Papers.

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

Ferrara, L., M. Mogliani, and J.-G. Sahuc (forthcoming): "High-frequency monitoring of growth-at-risk," *International Journal of Forecasting*.

Feunou, B., and R. Tédongap (2012): "A Stochastic Volatility Model With Conditional Skewness," *Journal of Business & Economic Statistics*, 30(4), 576–591.

Figueres, J. M., and M. Jarociński (2020): "Vulnerable growth in the euro area: Measuring the financial conditions," *Economics Letters*, 191(C).

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

Fissler, T., J. F. Ziegel, and T. Gneiting (2015): "Expected Shortfall is jointly elicitable with Value at Risk - Implications for backtesting," .

Frühwirth-Schnatter, S., and S. Pyne (2010): "Bayesian Inference for Finite Mixtures of Univariate and Multivariate Skew-Normal and Skew-t Distributions," *Biostatistics*, 11, 317—336.

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.

Gertler, M., and S. Gilchrist (2018): "What Happened: Financial Factors in the Great Recession," *Journal of Economic Perspectives*, 32(3), 3–30.

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

Giacomini, R., and H. White (2006): "Tests of Conditional Predictive Ability," *Econometrica*, 74(6), 1545–1578.

Gilchrist, S., and E. Zakrajšek (2012): "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, 102(4), 1692–1720.

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

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

Harvey, C. R., and A. Siddique (1999): "Autoregressive Conditional Skewness," *The Journal of Financial and Quantitative Analysis*, 34(4), 465–487.

He, Z., and A. Krishnamurthy (2012): "A Model of Capital and Crises," *Review of Economic Studies*, 79(2), 735–777.

——— (2013): "Intermediary Asset Prices," *American Economic Review*, 103(2), 1–43.

Hubrich, K., and R. J. Tetlow (2015): "Financial Stress and Economic Dynamics: The Transmission of Crises," *Journal of Monetary Economics*, 70, 100–115.

Iseringhausen, M. (2018): "The Time-Varying Asymmetry Of Exchange Rate Returns: A Stochastic Volatility – Stochastic Skewness Model," Working Papers of Faculty of Economics and Business Administration, Ghent University, Belgium 18/944, Ghent University, Faculty of Economics and Business Administration.

Jondeau, E., and M. Rockinger (2003): "Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements," *Journal of Economic Dynamics and Control*, 27(10), 1699–1737.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring Uncertainty," *American Economic Review*, 105(3), 1177–1216.

Justiniano, A., and G. E. Primiceri (2008): "The Time-varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 98(3), 601–641.

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

Kim, C.-J., J. Piger, and R. Startz (2008): "Estimation of Markov regime-switching regression models with endogenous switching," *Journal of Econometrics*, 143(2), 263 – 273.

Kiyotaki, N., and J. H. Moore (1997): "Credit Cycles," *Journal of Political Economy*, 105(2), 211–48.

Kremer, M., M. Lo Duca, and D. Holló (2012): "CISS - a composite indicator of systemic stress in the financial system," Working Paper Series 1426, European Central Bank.

Lhuissier, S. (2017): "Financial Intermediaries' Instability and Euro Area Macroeconomic Dynamics," *European Economic Review*, 98, 49 – 72.

——— (2018): "The Regime-Switching Volatility of Euro Area Business Cycles," *Macroeconomic Dynamics*, 22(2), 426—469.

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.

Lin, T. I., J. C. Lee, S. Y. Yen, and N. Chung (2007): "Finite mixture modelling using the skew normal distribution," *Statistica Sinica*.

Liu, Z., D. F. Waggoner, and T. Zha (2011): "Sources of Macroeconomic Fluctuations: A Regime-switching DSGE approach," *Quantitative Economics*, 2(2), 251–301.

Maggiori, M. (2017): "Financial Intermediation, International Risk Sharing, and Reserve Currencies," *American Economic Review*, 107(10), 3038–3071.

Nakajima, J. (2013): "Stochastic volatility model with regime-switching skewness in heavy-tailed errors for exchange rate returns," *Studies in Nonlinear Dynamics & Econometrics*, 17(5), 499–520.

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*.

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.

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.

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.

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

ONLINE APPENDIX: FINANCIAL CONDITIONS AND MACROECONOMIC DOWNSIDE RISKS IN THE EURO AREA

Not for Publication

STÉPHANE LHUISSIER

This Appendix consists of the following sections:

Appendix A. The Posterior

The conditional likelihood at time t is given by

$$p(y_t|Y_{t-1}, s_t, \theta), \tag{17}$$

which is generated by

$$p(y_t|Y_{t-1}, s_t, \theta) = \frac{2}{\sigma_{s_t^{scale}}}\phi\left(\frac{y_t - \mu_{s_t^{location}}}{\sigma_{s_t^{scale}}}\right)\Phi\left(\alpha_{s_t^{shape}}\frac{y_t - \mu_{s_t^{location}}}{\sigma_{s_t^{scale}}}\right). \tag{18}$$

Given (17), it follows that the overall likelihood of Y_T is

$$p(Y_T|\theta) = \prod_{t=1}^{T}\left[\sum_{s_t \in H} p(y_t|Y_{t-1}, s_t, \theta)\Pr(s_t, \theta)\right]. \tag{19}$$

The object inside the brackets of the likelihood in (19) can be interpreted as a weighted average of the conditional densities at time t given s_t. It can be evaluated recursively by updating $\Pr(s_t, \theta)$ according to the Hamilton (1989)'s filter. Interestingly, the inclusion of the additional shape parameter does not require to modify the original filter.

To form the posterior density, $p(\theta|Y_T)$, I combine the overall likelihood function $p(Y_T|\theta)$ with the prior $p(\theta)$:

$$p(\theta|Y_T) \propto p(Y_T|\theta)p(\theta), \tag{20}$$

The posterior density $p(\theta|Y_T)$ is not of standard form, but I show that it is possible to use the idea of Gibbs-sampling by sampling alternatively from conditional posterior distributions.

For computational reasons, I employ a logarithm transformation in equation (20) to obtain the log-posterior function as follows

$$\log\left\{p(\theta|Y_T)\right\} \propto \log\left\{p(Y_T|\theta)\right\} + \log\left\{p(\theta)\right\}, \tag{21}$$

where the conditional log-likelihood at time t, given s_t, is as follows

$$\begin{aligned}\log\left\{p(y_t|Y_{t-1}, s_t, \theta)\right\} =& \text{constant} - \log\left\{\sigma_{s_t^{scale}}\right\} \\ &- \frac{(y_t - \mu_{s_t^{location}})^2}{2\sigma^2_{s_t^{scale}}} + \log\left\{\Phi\left(\alpha_{s_t^{shape}}\frac{y_t - \mu_{s_t^{location}}}{\sigma_{s_t^{scale}}}\right)\right\}.\end{aligned}$$

Appendix B. Gibbs Sampler Procedure

Given the likelihood function and the prior density function, the objective is to obtain the conditional posterior density function by sampling alternately from the following conditional posterior distributions

(1) $p\left(S_T^{(n)}|Y_T, Z_T, \theta^{(n-1)}\right)$;
(2) $p\left(S_T^{*(n)}, \gamma^{(n)}|S_T^{(n)}, \Xi_T^{(n-1)}\right)$;
(3) $p\left(\Xi_T^{(n)}|Y_T, S_T^{(n)}, \theta^{(n-1)}\right)$;
(4) $p\left(\mu_\upsilon^{(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\mu}^{(n-1)}\right)$;
(5) $p\left(\sigma_\upsilon^{(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\sigma}^{(n)}\right)$;
(6) $p\left(\alpha_\upsilon^{(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\alpha}^{(n)}\right)$.

I now discuss each of these conditional density functions.

B.1. **Conditional posterior densities,** $p\left(S_T^{(n)}|Y_T, \theta^{(i-1)}\right)$**.** Following Filardo and Gordon (1998), the procedure for generating S_T is adapted from Albert and Chib (1993) approach for the fixed transition probability model. Because of the time-varying nature of transition probabilities, it is not anymore possible to generate S_T using the Carter and Kohn (1994)'s multi-move Gibbs-sampling procedure that allows to simulate S_T as a block. We need to modify the Albert and Chib (1993)'s single-move Gibbs-sampling procedure.

The full conditional distribution is

$$p(s_t^{(n)}|Y_T, Z_t, \theta^{(n-1)}) \propto p(s_t|s_{t-1}, z_t)p(s_{t+1}|s_t, z_{t+1}) \prod_{t=1}^{T} p(y_t|Y_{t-1}, S_t) \tag{22}$$

Drawing $S_T^{(n)}$ from the full conditional distribution based on this equation is as follows. I begin with a draw from $p(s_T|Y_T, \theta)$ obtained with the Hamilton (1989) basic filter, and then iterate recursively backward to draw $s_{T-1}, s_{T-2}, \ldots, 1$ using (22).

B.2. **Conditional posterior densities,** $p\left(S_T^{*(n)}, \gamma^{k(n)}|S_T, S_T^*, \Xi_T\right)$**.** Given values of γ^k and the inequality constraint in (5), values of S_T^* can be simulated from the appropriate truncated standard normal distributions.

Conditional on S_T, S_T^*, and Z_t, one can easily generate draws for γ^k using equation (6), which represents a simple linear regression model. Suppose that $W^{\mathrm{k}} = [\mathbb{1} \quad Z_T \quad S_{T-1}^{\mathrm{k}}]$. The conditional posterior distribution is then

$$p(\gamma^{k(n)}|S_T^{\mathrm{k}}, S_T^{\mathrm{k}*}, Z_T) = \text{normal}(m_\gamma^{\mathrm{k}}, M_\gamma^{\mathrm{k}}), \tag{23}$$

where $M_\gamma^{\mathrm{k}} = (\bar{\gamma}_2^{-1} + W^{\mathrm{k}\prime}W^{\mathrm{k}})^{-1}$, and $m_\gamma^{\mathrm{k}} = M_\gamma^{\mathrm{k}}(\bar{\gamma}_2^{-1}\bar{\gamma}_1 + W^{\mathrm{k}\prime}S_T^{\mathrm{k}*})$.

B.3. **Conditional posterior densities,** $p\left(\Xi_T^{(n)}|Y_T, S_T^{(n)}, \theta^{(i-1)}\right)$**.** Here, the nice property of such a model is that the full conditional distribution of Ξ_t given Y_t, $S_T^{(n)}$, and $\theta^{(n)}$ is available in closed form.

For $t = 1, 2, ..., T$, I generate $\Xi_T^{(n)}$ according to

$$p\left(\Xi_T^{(n)}|Y_T, S_T^{(n)}, \theta^{(i-1)}\right) = \prod_{t=1}^{T} p\left(\xi_t^{(n)}|Y_t, S_t^{(n)}, \theta^{(i-1)}\right), \tag{24}$$

where

$$p\left(\xi_t^{(n)}|Y_t, S_t^{(n)}, \theta^{(i-1)}\right) = \text{truncated-normal}\left(\xi_t^{(n)}|\delta_{s_t}^{(n)}(y_t - \mu_{s_t}^{(n)}), \sigma_{s_t}^2(1 - \delta_{s_t}^{(n)2})\right)_{\xi_t^{(n)}>0}. \tag{25}$$

B.4. **Conditional posterior densities,** $p\left(\mu_v^{(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\mu}^{(i-1)}\right)$**.** If we let $y_t^* = \frac{y_t - \delta_{s_t}\xi_t}{\sqrt{1-\delta_{s_t}^2}}$, and $x_t^* = \frac{x_t}{\sqrt{1-\delta_{s_t}^2}}$, we obtain an homoskedastic model as follows

$$y_t^* = \phi x_t^* + \nu_t, \tag{26}$$

where ν_t follows a standard normal distribution. Then, simulation from the full conditional distribution of $\phi^{(n)}$, given Y_T, $S_T^{(n)}$, $\Xi_T^{(n)}$ and $\theta_{\neq\phi}^{(i-1)}$, becomes straightforward, given a conjugate prior distribution. The posterior is defined as

$$p\left(\phi^{(n)}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\phi}^{(i-1)}\right) = \text{normal}\left(m_\mu^{(n)}, M_\mu^{(n)}\right), \tag{27}$$

where

$$m_\mu^{(n)} = \left(\bar{\mu}_2^{-1} + X'X\right)^{-1}\left(\bar{\mu}_2^{-1}\bar{\mu}_1 + X'y_t^*\right), \tag{28}$$

$$M_\mu^{(n)} = \left(\bar{\mu}_2^{-1} + X'X\right)^{-1}, \tag{29}$$

and $\bar{\mu}_1$ and $\bar{\mu}_2$ are known hyperparameters of the prior distribution — the mean and the variance, respectively — and $X = [x_1^*, \ldots, x_T^*]'$.

B.5. **Conditional posterior densities,** $p\left(\frac{1}{\sigma_v^{2(n)}}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\sigma}^{(n)}\right)$**.** Given Y_t, S_T, Ξ_T, θ, and S_T, the scale parameter $\frac{1}{\sigma^2}$ can be drawn using the following gamma distribution

$$p\left(\frac{1}{\sigma_v^{2(n)}}|Y_T, S_T^{(n)}, \Xi_T^{(n)}, \theta_{\neq\sigma}^{(n)}\right) = \text{gamma}(\tilde{\alpha}, \tilde{\beta}), \tag{30}$$

where

$$\tilde{\alpha} = \bar{\alpha} + T_v,$$

$$\tilde{\beta} = \bar{\beta} + \frac{1}{2\left(1 - \delta_{s_t}^{2\ (n)}\right)} \sum_{t\in\{t:s_t=v\}} \left(\left(y_t - \mu_{s_t}^{(n)}\right)^2 - 2\left(y_t - \mu_{s_t}^{(n)}\right)\delta_{s_t}^{(n)}\xi_t^{(n)} + \xi_t^{2(n)}\right),$$

with T_υ is the number of elements in $\{t : s_t^{\text{scale}} = \upsilon\}$ for $\upsilon = 1, 2$, and $\bar{\alpha}$ and $\bar{\beta}$ are the hyperparameters.

B.6. **Conditional posterior densities,** $p\left(\alpha_\upsilon^{(n)}|Y_T, S_T^{(n)}, \theta_{\neq\alpha}^{(n)}\right)$**.** Let $\bar{y}_t = \frac{y_t - \mu_{s_t}}{\sigma_{s_t}}$ and $\overline{Y}_T = [\bar{y}_1, \ldots, \bar{y}_T]'$. Consider the following derivation for the full conditional distribution of α_υ, given Y_T, $S_T^{(n)}$, and $\theta_{\neq\alpha}^{(n)}$:

$$\begin{aligned}
p(\alpha_\upsilon^{(n)}|Y_T, S_T^{(n)}, \theta_{\neq\alpha}^{(n)}) &\propto \phi\left(\frac{\alpha_\upsilon - \alpha_0}{\psi_0}\right) \prod_{t=1}^{T} \Phi\left(\alpha_\upsilon \bar{y}_t\right) \\
&\propto \phi\left(\frac{\alpha_\upsilon - \alpha_0}{\psi_0}\right) \Phi_T\left(\alpha_\upsilon \overline{Y}_T; I_T\right) \\
&\propto \phi\left(\frac{\alpha_\upsilon - \alpha_0}{\psi_0}\right) \Phi_T\left(\overline{Y}_T \alpha_0 + \overline{Y}_T(\alpha_\upsilon - \alpha_0); I_T\right) \\
&\propto \text{SUN}_{1,T}\left(\alpha_\upsilon^{(n)}|\alpha_0, \Delta_1(k)\alpha_0/\psi_0, \psi_0, 1, \Delta_1, \Gamma_1\right)
\end{aligned}$$

where $\text{SUN}_{d,m}(x|\xi, \tau, \omega, \Omega, \Delta, \Gamma)$ refers to the unified skew-normal (SUN) distribution introduced by Arellano-Valle and Azzalini (2006) as follows

$$\phi_d\left(z - \xi; \omega\Omega\omega\right) \frac{\Phi_m\left(\gamma + \Delta\Omega^{-1}\omega^{-1}(z - \xi); \Gamma - \Delta\Omega^{-1}\Delta'\right)}{\Phi_m(\gamma; \Gamma)^{-1}}, \tag{31}$$

with Φ_d is the cumulative density function of d-variate Gaussian distribution with variance-covariance matrix Σ, Ω, Γ, and $\Omega^* = ((\Gamma, \Delta)', (\Delta', \Omega)')$ are correlations matrices, and ω is a $d \times d$ diagonal matrix; $\Delta_1 = [\zeta_t]_{t=1,\ldots,T}$ with $\zeta_t = \psi_0 \bar{y}_t^2 (\psi_0^2 \bar{y}_t^2 + 1)^{-1/2}$; $\Gamma_1 = I - \text{diag}(\Delta_1)^2 + \Delta_1 \Delta_1^2$; and where $\text{diag}(V)$ is a diagonal matrix, the elements of which coincide with those of vector V.[15]

To simulate draws from the SUN distribution, one can use its stochastic representation. Let U_0 and $U_{1,-\gamma}$ have the following distribution

$$U_0 = \text{normal}\left(U_0|0, \overline{\Psi}_\Delta\right), \quad \text{and} \quad U_{1,-\gamma} = \text{truncated-normal}(U_1|0, \Gamma)_{-\gamma}. \tag{32}$$

Then, it can be show the SUN distribution can be generated as follows

$$\xi + \omega\left(U_0 + \Delta\Gamma^{-1} U_{1,-\gamma}\right). \tag{33}$$

Once we obtain α_υ, we can directly transform it to recover $\delta_\upsilon = \frac{\alpha_\upsilon}{\sqrt{(1+\alpha_\upsilon^2)}}$.

[15] Canale, Pagui, and Scarpa (2016) demonstrates that informative priors (i.e., normal or skew-normal distribution) for the shape parameter of a constant skew-normal model lead to closed-form full conditional distributions.

Appendix C. The Moments

Deriving the first four centered, conditional moments of the Markov-switching model is not straightforward. Recall that $p_{i,t} = \Pr(s_t = i|Y_{t-1}, \theta)$ is the probability of being in regime i, with $i \in \{1, \ldots, H\}$, at time t given information $t-1$. Proposition 1 in Timmermann (2000) characterizes the moments of the basic Markov switching model from the law of iterated expectations as follows

$$\mathrm{E}[(y_t - \tilde{\mu}_t)^n|Y_{t-1}, \theta] = \mathrm{E}\left[\mathrm{E}[(y_t - \tilde{\mu}_t)^n|Y_{t-1}, \theta, S_t]\right] \tag{34}$$

$$= \sum_{i=1}^{H} \mathrm{E}[(\mu_i + \sigma_i \varepsilon_t - \tilde{\mu}_t)^n|Y_{t-1}, \theta] \tag{35}$$

Initially, Timmermann (2000) develops the expressions for the cases where ε_t follows a t-distribution or a normal distribution. I extend the approach for the case where ε_t is a skew-normal distribution. It implies

$$\mathrm{E}[(y_t - \tilde{\mu}_t)^n|Y_{t-1}, \theta] = \sum_{i=1}^{H} p_{i,t} \sum_{j=0}^{n} C_j^n \sigma_i^j \mathrm{E}(\varepsilon_t^j)(\mu_i - \tilde{\mu}_t)^{n-j}, \tag{36}$$

with $C_j^n = \frac{n!}{(n-j)!j!}$. More specifically, the centered moments are as follows

$$\mathrm{E}[(y_t - \tilde{\mu}_t)^n|Y_{t-1}, \theta] = \sum_{i=1}^{H} p_{i,t} \sum_{j=0}^{n} C_j^n \sigma_i^j a_{i,j}(\mu_i - \tilde{\mu}_t)^{n-j}, \tag{37}$$

where

$$a_{i,j} = \frac{(\alpha_i\sqrt{2})^n}{\sqrt{\pi}\alpha_i^n} C_j^n \left(\frac{1}{\alpha_i\sqrt{2}}\right)^j \Gamma\left(\frac{n-j-1}{2}\right) b_j, \tag{38}$$

with $\Gamma(.)$ is the gamma distribution and $b_j = \mathrm{E}(V^j) = \begin{cases} 0, & \text{if } j \text{ odd} \\ 2^{j/2}\Gamma\left(\frac{j+1}{2}\right)/\sqrt{\pi} & \text{if } j \text{ even} \end{cases}$

While the first and second moments, $\mathrm{E}[y_t|Y_{t-1}, \theta] = \tilde{\mu}_t$ and $\mathrm{E}[(y_t - \tilde{\mu}_t)^2|Y_{t-1}, \theta]$ 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

$$\sqrt{b_1} \equiv \frac{\mathrm{E}[(y_t - \mu_t)^3]}{\left(\mathrm{E}[(y_t - \mu_t)^2]\right)^{3/2}}, \qquad b_2 \equiv \frac{\mathrm{E}[(y_t - \mu_t)^4] - 3\left(\mathrm{E}[(y_t - \mu_t)^2]\right)^2}{\left(\mathrm{E}[(y_t - \mu_t)^2]\right)^2}. \tag{39}$$

Appendix D. Forecasting

I focus on the skew-normal model with Markov-switching. I generate draws from the posterior predictive density using the following decomposition

$$p(Y_{T+1:T+h}|Y_T) = \int_{(\theta,S_T)} \left[\int_{S_{T+1:T+h}} p(Y_{T+1:T+h}|S_{T+1:T+h}) \times p(S_{T+1:T+h}|\theta, S_T, Y_T) d(S_{T+1:T+h}) \right] \times p(\theta, S_T|Y_T) d(\theta, S_T). \quad (40)$$

I use the sequence $t_1 : t_2$ to indicate from t_1 to t_2.The decomposition shows how the predictive density reflects about parameters and regimes at the forecast origin, $p(\theta, S_T|Y_T)$, and uncertainty about future regimes. Motivated by this decomposition, I generate draws from the predictive density taking into account the hidden Markov regimes s_t.

Algorithm 1 *Predictive Density Draws.*

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

(1) Draw $\left(\theta^{(j)}, S_T^{(j)}\right)$ from the posterior distribution $(\theta, S_T|Y_T)$.

(2) Draw $p\left(S_{T+1:T+h}|\theta^{(j)}, S_T^{(j)}\right)$ as follows:

(a) Compute the transition matrix $Q_T^{(j)}$ using the information z_T.

(b) Draw the sequence of unobservable regimes $S_{T+1:T+h}^{(j)}$ from the transition matrix $Q_T^{(j)}$.

(3) Compute the sequence $Y_{T+1:T+h}^{(j)}$ by drawing from the following distribution

$$\text{skew-normal}\left(Y_{T+1:T+h}^{(j)}|\mu_{s_t^{(j)}\text{location}}^{(j)}, \sigma^{2(j)}_{s_t^{(j)}\text{scale}}, \alpha_{s_t^{(j)}\text{shape}}^{(j)}\right), \qquad t = T+1, \ldots, H. \quad (41)$$

Algorithm 1 produces n_{sim} trajectories $Y_{T+1:T+h}^{(j)}$ from the predictive distribution of $Y_{T+1:T+h}^{(j)}$ given Y_T. In Step 1, I consider the Gibbs sampling to obtain the empirical joint posterior density $\left(\theta^{(j)}, S_T^{(j)}\right)$. I generate $N_1 + N_2 = 55,000$ draws, the first $N_1 = 5,000$ are discarded as burn-in and of the remaining $N_2 = 50,000$ draws, one of every 10 draws is retained to get $5,000$ draws of parameters and sequences of regimes. In Step 2, it is crucial to generate future paths of the unobservable state s_t to iterate forward GDP growth. In Step 3, I draw from the skew-normal distribution conditional on the sequence of regimes.

When computing moments of the one-step-ahead out-of-sample forecast distribution in Section VI, I proceed as follows. For each parameter draws, and for each point in time, I simulate $1,000$ one-quarter-ahead paths following Steps 2 and 3. I then compute various moments of the distribution of GDP growth.

Appendix E. Robustness

E.1. **Prior Sensitivity.** Tables 5 and 6 report prior and posterior distributions for alternative prior specifications for γ_0 and γ_1. Figures 9 and 10 report the resulting moments and downside risks, respectively.

Table 5. Prior and Posterior Distributions.

	Prior			Posterior					
Coefficient	Density	para(1)	para(2)	Mean	Median	[16;	84]	[5;	95]
Skew-normal parameters									
$\mu(s^{\text{location}}=1)$	N	0.00	2.00	−0.0116	−0.0242	−0.1413	0.1440	−0.2256	0.2513
$\mu(s^{\text{location}}=2)$	N	0.00	2.00	0.5593	0.5614	0.5204	0.6005	0.4766	0.6355
$1/\sigma^2(s^{\text{scale}}=1)$	G	1.00	1.00	0.3741	0.3249	0.1541	0.5943	0.0812	0.8225
$1/\sigma^2(s^{\text{scale}}=2)$	G	1.00	1.00	8.3036	8.2076	6.7056	9.9305	5.7691	11.1105
$\alpha(s^{\text{shape}}=1)$	N	0.00	4.00	−4.3589	−3.9792	−6.3566	−2.3740	−8.3341	−1.6837
$\alpha(s^{\text{shape}}=2)$	N	0.00	4.00	4.0179	3.6447	1.7155	6.3593	0.8358	8.5257
Transition probability parameters, s^{location}									
$\gamma_0^{\text{location}}$	N	−0.70	0.25	−0.6288	−0.6254	−0.8508	−0.4085	−0.9992	−0.2641
$\gamma_1^{\text{location}}$	N	2.20	0.25	2.2894	2.2892	2.0545	2.5192	1.9046	2.6835
$\gamma_{\text{gdp},1}^{\text{location}}$	N	0.00	2.00	−0.0752	−0.1297	−1.0581	0.9319	−1.7182	1.7651
$\gamma_{\text{ciss},1}^{\text{location}}$	N	0.00	2.00	−0.9616	−0.9414	−2.8181	0.9084	−4.0988	2.1228
$\gamma_{\text{gdp},2}^{\text{location}}$	N	0.00	2.00	0.7672	0.6727	−0.7859	2.3478	−1.5868	3.4765
$\gamma_{\text{ciss},2}^{\text{location}}$	N	0.00	2.00	−1.7466	−1.8004	−3.5314	0.0525	−4.6956	1.2996
Transition probability parameters, s^{scale}									
γ_0^{scale}	N	−0.60	0.25	−0.4128	−0.4177	−0.6248	−0.1955	−0.7595	−0.0520
γ_1^{scale}	N	2.30	0.25	2.4562	2.4553	2.2330	2.6784	2.0941	2.8295
$\gamma_{\text{gdp},1}^{\text{scale}}$	N	0.00	2.00	0.6350	0.3624	−0.6216	1.9694	−1.1005	3.1836
$\gamma_{\text{ciss},1}^{\text{scale}}$	N	0.00	2.00	0.0972	0.1216	−1.8962	2.1174	−3.2999	3.3016
$\gamma_{\text{gdp},2}^{\text{scale}}$	N	0.00	2.00	0.9071	0.9591	−0.1106	1.8769	−0.8836	2.5869
$\gamma_{\text{ciss},2}^{\text{scale}}$	N	0.00	2.00	−0.9818	−0.9699	−2.6365	0.7056	−3.7608	1.7166
Transition probability parameters, s^{shape}									
γ_0^{shape}	N	−1.00	0.25	−1.1020	−1.1055	−1.2959	−0.9044	−1.4209	−0.7737
γ_1^{shape}	N	2.00	0.25	1.9673	1.9710	1.7313	2.2009	1.5748	2.3551
$\gamma_{\text{gdp},1}^{\text{shape}}$	N	0.00	2.00	1.8965	1.7639	0.5720	3.2478	−0.0774	4.3174
$\gamma_{\text{ciss},1}^{\text{shape}}$	N	0.00	2.00	0.1348	0.1687	−1.4995	1.7720	−2.5902	2.7670
$\gamma_{\text{gdp},2}^{\text{shape}}$	N	0.00	2.00	0.1313	−0.0287	−0.9119	1.1112	−1.4895	2.2875
$\gamma_{\text{ciss},2}^{\text{shape}}$	N	0.00	2.00	−0.6572	−0.6584	−2.5898	1.2279	−3.8008	2.5754

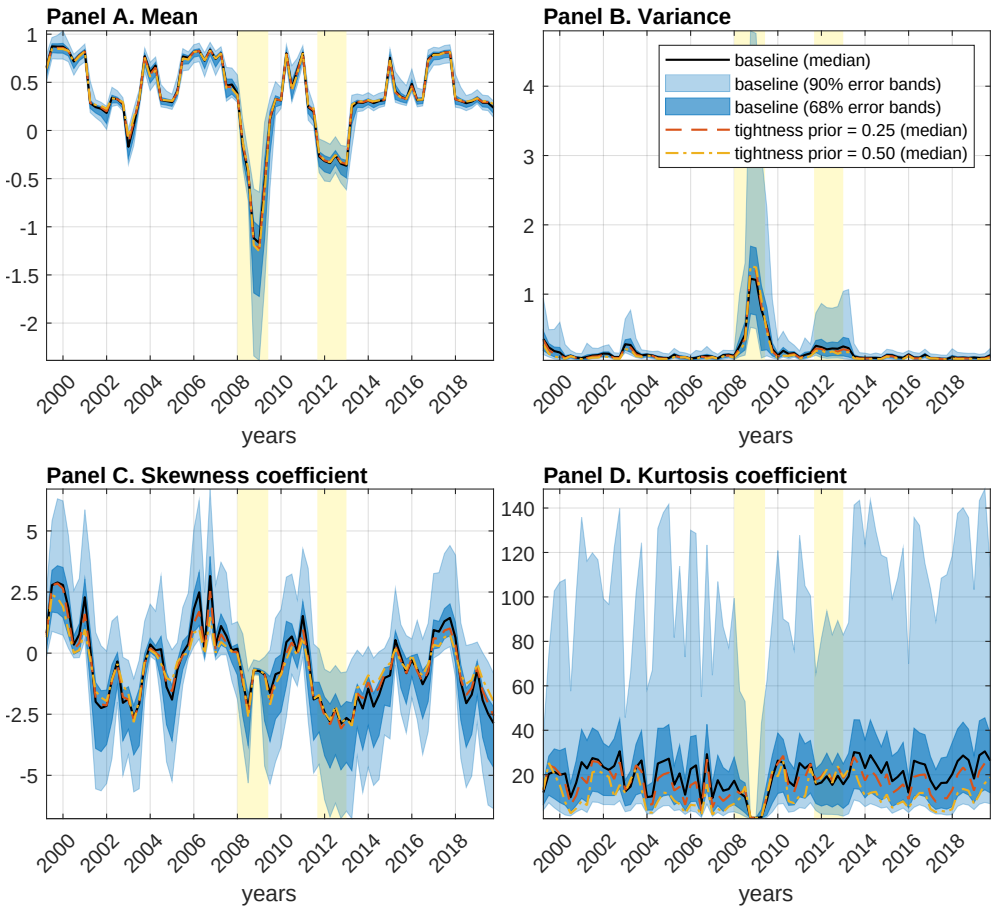
Note: N stands for Normal, and G for Gamma distributions. Para(1) and Para(2) correspond to the means and standard deviations for Normal Distributions, and to the hyperparameters for Gamma Distributions.

TABLE 6. Prior and Posterior Distributions.

	Prior			Posterior					
Coefficient	Density	para(1)	para(2)	Mean	Median	[16;	84]	[5;	95]
Skew-normal parameters									
$\mu(s^{\text{location}}=1)$	N	0.00	2.00	−0.0407	−0.0511	−0.1557	0.1124	−0.2477	0.2177
$\mu(s^{\text{location}}=2)$	N	0.00	2.00	0.5683	0.5659	0.5277	0.6063	0.4922	0.6462
$1/\sigma^2(s^{\text{scale}}=1)$	G	1.00	1.00	0.3673	0.3141	0.1487	0.5787	0.0883	0.8254
$1/\sigma^2(s^{\text{scale}}=2)$	G	1.00	1.00	8.2657	8.1739	6.5953	9.9024	5.6683	11.1517
$\alpha(s^{\text{shape}}=1)$	N	0.00	4.00	−4.2302	−3.8225	−6.2348	−2.2763	−8.2749	−1.5506
$\alpha(s^{\text{shape}}=2)$	N	0.00	4.00	3.9009	3.5195	1.5985	6.3115	0.6850	8.6000
Transition probability parameters, s^{location}									
$\gamma_0^{\text{location}}$	N	−0.70	0.50	−0.5800	−0.5900	−0.9824	−0.1654	−1.2414	0.0967
$\gamma_1^{\text{location}}$	N	2.20	0.50	2.4183	2.4199	2.0002	2.8398	1.7325	3.1313
$\gamma_{\text{gdp},1}^{\text{location}}$	N	0.00	2.00	−0.0388	−0.1156	−1.1160	1.1316	−1.8076	2.0182
$\gamma_{\text{ciss},1}^{\text{location}}$	N	0.00	2.00	−0.8820	−0.8674	−2.7658	1.0120	−4.0182	2.2061
$\gamma_{\text{gdp},2}^{\text{location}}$	N	0.00	2.00	0.8404	0.7608	−0.8890	2.5586	−1.7665	3.7225
$\gamma_{\text{ciss},2}^{\text{location}}$	N	0.00	2.00	−1.8136	−1.8664	−3.6475	−0.0059	−4.7483	1.3444
Transition probability parameters, s^{scale}									
γ_0^{scale}	N	−0.60	0.50	−0.2161	−0.2109	−0.6217	0.1791	−0.8713	0.4518
γ_1^{scale}	N	2.30	0.50	2.5998	2.6025	2.1881	3.0124	1.9097	3.2775
$\gamma_{\text{gdp},1}^{\text{scale}}$	N	0.00	2.00	0.7986	0.6330	−0.5325	2.1906	−1.0305	3.1877
$\gamma_{\text{ciss},1}^{\text{scale}}$	N	0.00	2.00	0.0739	0.1239	−1.9737	2.1226	−3.3342	3.2735
$\gamma_{\text{gdp},2}^{\text{scale}}$	N	0.00	2.00	1.1572	1.2290	−0.0029	2.3240	−1.0013	3.0327
$\gamma_{\text{ciss},2}^{\text{scale}}$	N	0.00	2.00	−1.0282	−1.0216	−2.8197	0.7516	−3.8923	1.8724
Transition probability parameters, s^{shape}									
γ_0^{shape}	N	−1.00	0.50	−1.1913	−1.1813	−1.4633	−0.9172	−1.6920	−0.7368
γ_1^{shape}	N	2.00	0.50	1.9336	1.9264	1.5324	2.3329	1.2592	2.6272
$\gamma_{\text{gdp},1}^{\text{shape}}$	N	0.00	2.00	2.1071	2.0008	0.6756	3.5685	−0.0085	4.5995
$\gamma_{\text{ciss},1}^{\text{shape}}$	N	0.00	2.00	0.1423	0.1446	−1.4712	1.7950	−2.5636	2.8152
$\gamma_{\text{gdp},2}^{\text{shape}}$	N	0.00	2.00	0.2123	0.1197	−0.9167	1.3099	−1.5703	2.3501
$\gamma_{\text{ciss},2}^{\text{shape}}$	N	0.00	2.00	−0.7061	−0.7152	−2.6479	1.1942	−3.7954	2.5175

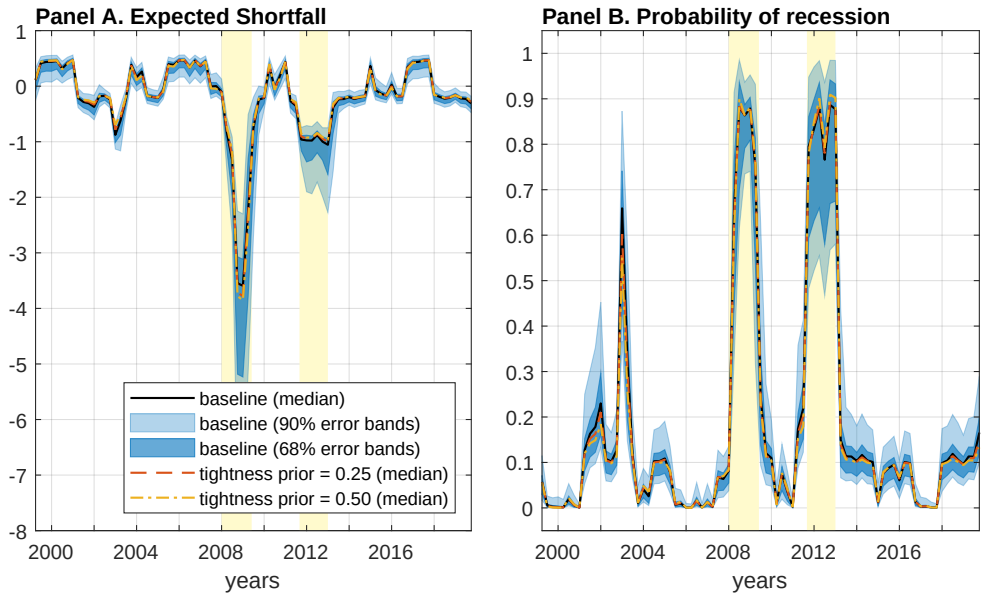
Note: N stands for Normal, and G for Gamma distributions. Para(1) and Para(2) correspond to the means and standard deviations for Normal Distributions, and to the hyperparameters for Gamma Distributions.

FIGURE 9. Moments with alternative prior specifications.



Note: Sample period: 1999.Q2 — 2019.Q4. Historical path of first four moments — mean, variance, skewness, and kurtosis — of euro area GDP growth. The yellow areas denote the CEPR recessions.

FIGURE 10. Expected shortfall and probability of recession with alternative prior specifications.



Note: Sample period: 1999.Q2 — 2019.Q4. Evolution of the 5% expected shortfall and the probability of recession. The yellow areas denote the CEPR recessions.

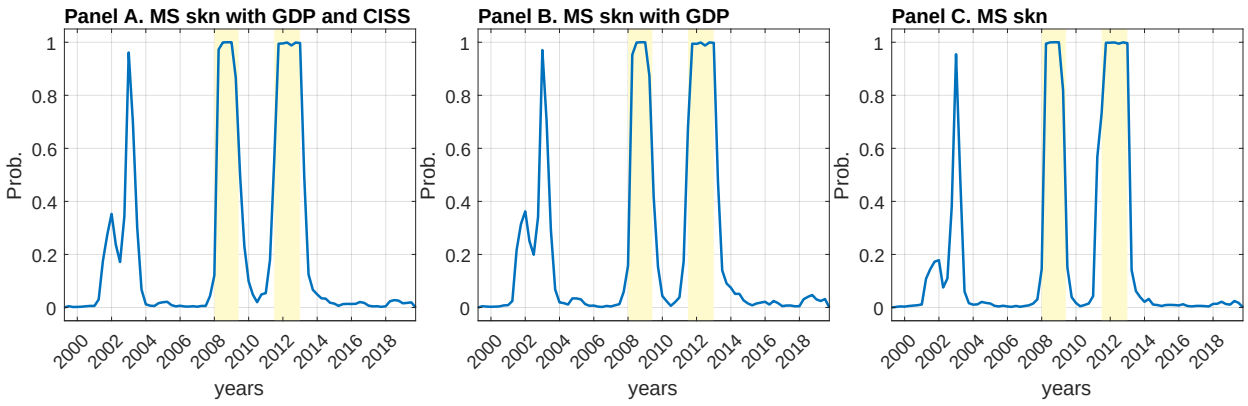
E.2. **Independent-chains model vs synchronized-chains model.** Table 7 compares the independent-chains model and the synchronized-chains model using an information criteria. Figure 11 shows the probabilities of being in Regime 1 under the synchronized-chains models.

TABLE 7. Information criteria

	Difference relative to synchronized-chains model	
Independent-chains model	WAIC	standard errors
MS skew-normal with GDP + CISS	−11.1786	4.4467
MS skew-normal with GDP	−11.5745	4.5779
MS skew-normal	−14.4201	5.3460

Note: Watanabe-Akaike information criteria (WAIC). WAIC $= -\log(\overline{lpd}) + \overline{p}$, where $\log(\overline{lpd})$ is the log pointwise predictive density, i.e., $\sum_{t=1}^{T} \log\left(\frac{1}{S}\sum_{s=1}^{S} p(y_i|\theta^s)\right)$ with S is the number of MCMC iteration, and $\overline{p}$ is the estimated effective number of parameters, i.e., $\sum_{t=1}^{T} V_{s=1}^{S}(\log(p(y_i|\theta^s))$, with V represent the sample variance. The standard error $se(\overline{elpd}) = \sqrt{(T * V_{t=1}^{T}\overline{elpd,t})}$, where $\overline{elpd,t} = \log\left(\frac{1}{S}\sum_{s=1}^{S} p(y_i|\theta^s)\right) - \left(V_{s=1}^{S}\log p(y_i|\theta^s)\right)$. For model comparison between A and B, the standard error is $se(\overline{elpd}^A - \overline{elpd}^B) = \sqrt{T * V_{t=1}^{T}(\overline{elpd,t}^A - \overline{elpd,t}^B)}$. Negative values of WAIC differences indicate that the independent-chains model performs better than the synchronized-chains model.

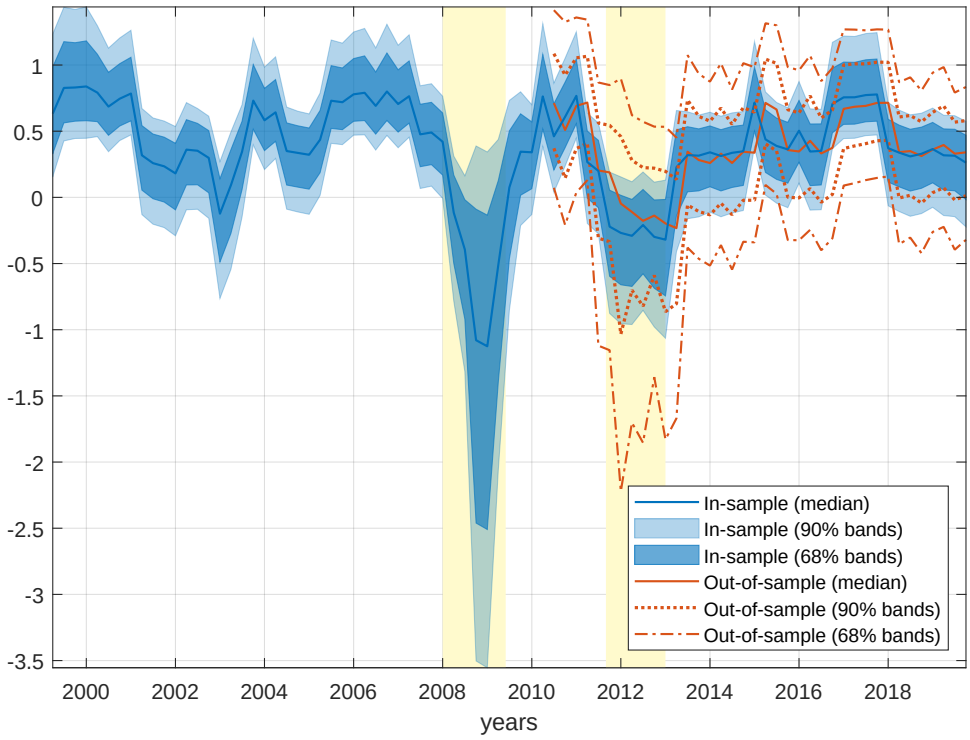
FIGURE 11. Regime probabilities.



Note: Sample period: 1999.Q2 — 2019.Q4. Probabilities of being in Regime 1 produced from the synchronized models: MS skew-normal with GDP and CISS (Panel A), MS skew-normal with GDP (Panel B) and MS skew-normal (Panel C). The location, scale, and shape parameters follow the same two-state Markov-switching process. The yellow areas denote the CEPR recessions.

E.3. **In-sample vs out-of-sample forecasts.** Figure 12 compares out-of-sample and in-sample predictive densities for GDP growth one quarter ahead using the baseline model (MS skew-normal with GDP and CISS). Table 8 reports the accuracy of in-sample forecasts of GDP growth using three forecast metrics : MSFE, log-score, and tail-risk scores. For the latter, I consider three basic measures: (1) the percentage of outcomes falling below the 5 percent quantile of the forecast distribution; (2) the quantile score; and (3) the joint value at risk-expected shortfall (VaR-ES) score.

FIGURE 12. Out-of-sample vs in-sample predictions



Note: Sample period: 1999.Q2 — 2019.Q4. The figure compares out-of-sample and in-sample predictive densities for GDP growth one quarter ahead.

TABLE 8. Accuracy of in-sample forecasts of GDP growth

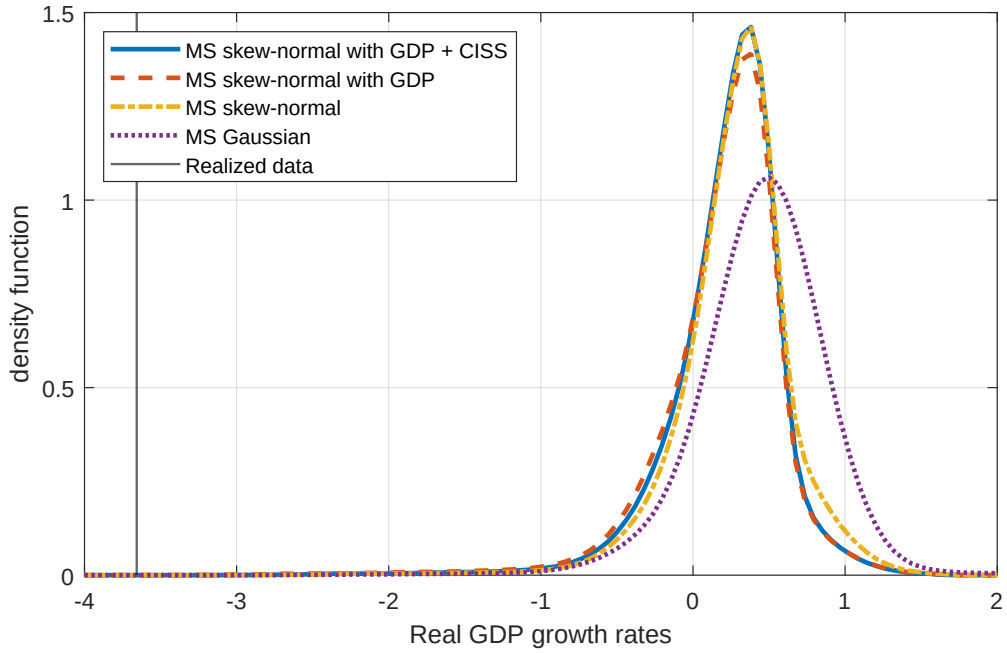
Model	MSFE	Log-Score	Tail risk scores		
			5% coverage	Quantile	VaR-ES
MS skew-normal with GDP and CISS	0.5373* (0.0001)	0.4268* (0.0000)	0.0068* (0.0000)	0.7757* (0.0014)	−0.0997* (0.0003)
MS skew-normal with GDP	0.5443* (0.0001)	0.4148* (0.0000)	0.0067* (0.0000)	0.7730* (0.0010)	−0.0989* (0.0002)
MS skew-normal	0.4935* (0.0000)	0.4605* (0.0000)	0.0049* (0.0000)	0.6848* (0.0000)	−0.1197* (0.0000)

Note: The table reports average forecast metrics relative to the MS Gaussian model, except in the case of the 5 percent coverage rates. I use ratios for the MSFE, Quantile, and differences for Log-Score and VaR-ES. Ratios smaller than 1, and positive values of the log-score differences indicate that the model performs better than the MS Gaussian benchmark. Values in parentheses report the p-values of the Diebold and Mariano (1995)-West (1996) test statistic for equal predictive accuracy. "*" denotes significance at the 5% level. Significance of the test follows from Giacomini and White (2006)'s critical values.

Appendix F. Covid-19 Crisis

Figure 13 shows the predictive densities of GDP growth for the first quarter of 2020.

Figure 13. Predictive Densities of GDP Growth in the COVID-19 Crisis



Note: Predictive densities produced from the models under consideration: MS skew-normal with GDP + CISS (baseline model), MS skew-normal with GDP, MS skew-normal, and MS Gaussian.