THE SWITCHING SKEWNESS OVER THE BUSINESS CYCLE

STÉPHANE LHUISSIER*

Abstract. Motivated by characterizing the cyclical variations in the upside and downside risks of GDP growth, we examine autoregressive time series models subject to Markov skewness shifts within a Bayesian framework. Our methodology leads to closed-form full conditional posterior distributions, whose sampling can be efficiently conducted within a Gibbs sampling scheme. We estimate the model for postwar U.S. GDP growth and reveal that time variations in GDP skewness is a recurrent feature of the U.S. business cycle. Consistent with macro-finance theory, incorporating additional information from financial markets help to identify a regime marked by downside tail risk to growth.

I. Introduction

How does the distribution of the growth rate of macroeconomic time series data evolve across various phases of the "business cycle"? Since the first half of the past century, this key question about business cycle asymmetry has been the focus of a large part of macroeconomics. All major works on business cycles converge toward the same conclusion that the dynamics of the first two moments — the mean and the variance — are asymmetric over different phases of the cycle. For example, recession phase is commonly associated with negative and volatile GDP growth rates, while expansion phase is characterized by positive and stable GDP growth rates. However, little is known about asymmetries in other than the first two moments.

Business cycle asymmetries of GDP growth need to be addressed in the context of non-Gaussian density as emphasized by macroeconomic risk models. For example, Rietz (1988)

Date: January 11, 2019.

Key words and phrases. regime switching, skewness, business cycle, Gibbs-sampler, macro-finance, upside and downside risk.

JEL Classification: C11; C2; E32.

*Banque de France, 31, Rue Croix des Petits Champs, DGSEI-DEMFI-POMONE 41-1422, 75049 Paris Cedex 01, FRANCE (Email: stephane_lhuisier@club.fr; URL: http://www.stephanelhuissier.eu). The views expressed in this paper are those of the author and should under no circumstances be interpreted as reflecting those of the Banque de France or the Eurosystem. I thank Tobias Adrian, Isaac Baley, Marco Del Negro, Kyle Jurado, Jean-Marc Robin, Moritz Schularick, Mathias Trabandt and participants at the 10th French Econometrics Conference (PSE), and the 12th International Conference on Computational and Financial Econometrics (University of Pisa) for their helpful comments.

characterizes risk through the probability that an economic disaster (i.e., a negative economic shock) occurs in order to explain a lot of puzzles about asset returns. Similarly, Barro (2006, 2009) argue that the possibility of rare economic disasters is a substantial driver for equity premium in asset markets. Gourio (2012) augments a standard dynamic stochastic general equilibrium (DSGE) model to include a small risk of economic disaster and argues that an increase in the probability of a disaster generates large contraction in spending, investment and output. Additionally, the notion of downside and upside risks to GDP growth has received a great deal of attention from policymakers over the past few years. For example, Yellen (2016) highlights how downside risks to the outlook for economic activity may affect the Federal Open Market Committee (FOMC)'s policy deliberations.

In this paper we focus on the autoregressive (AR) model that is subject to third-moment (skewness) shifts, thus making allowance for asymmetric changes in the mass of the distribution over time.[1] Specifically, the skewness of the time series is governed by a Markov-switching process with unknown transition probabilities. Hamilton (1989) uses a similar process to capture changes in the mean of the growth rates of postwar U.S. real Gross National Product (GNP) between recession periods and expansion periods. Since then, it has been intensively demonstrated that this tool has been successful in capturing regime changes in a variety of areas of macroeconomics. For a recent survey of this literature see Hamilton (2016). Our approach here is to propose a simple and easy-to-implement Bayesian framework for studying the behavior of skewness of macroeconomic time series data over time. The starting point of the paper is to consider a skew-normal distribution proposed by Azzalini (1985, 1986). The key innovation in his work is to allow for possible departure from symmetry to produce more realistic and more flexible families of distributions to describe data. With respect to the normal distribution, the skew-normal family is a class of density functions that depends on an additional shape parameter that affects the tails of the density. Such a distribution provides a characterization of variations in risk that goes beyond what can be achieved through the typical Gaussian assumption.

The main contribution of the paper is twofold. First, we develop a Gibbs sampler for Bayesian inferences of AR time series subject to Markov skewness shifts. Our Gibbs sampling procedure can be seen as an extension of Albert and Chib (1993) for univariate models to regime-switching skewness. Specifically, we 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 straightfward Markov Chain Monte Carlo (MCMC) sampling sequence that involves a 6-block Gibbs sampler for Markov-switching models of conditional

[1]The skewness measures the asymmetry of a given distribution. A negative skewness means that negative outcomes are more likely than bad outcomes of the same magnitude.

skewness, in which one can generate in a flexible and straightforward manner alternatively draws from full conditional posterior distributions. In order to make computationally feasible estimation and inference of AR model with Markov skewness shifts, we provide a companion software package for anyone interested in such models. We apply our methodology to the growth rates of postwar U.S. real GDP, and show that the distribution of GDP growth exhibits large variation over time, where changes in tail probabilities are unambiguously strong. The times of changes are remarkably similar to NBER dating of business cycles. In particular, we find that expansion times are characterized by large positive skewness — and, so, a large right tail of positive GDP growth rates. That is, we observe an increase in the probability of very large positive shocks in expansion periods (i.e., upside risk). By contrast, during recession periods, the distribution of GDP growth is close to being symmetric. The results are robusts when taking into account the possibility of changes in the mean and in the variance of GDP growth.

Second, we test the nonlinear relationship between GDP skewness and financial conditions, as emphasized by the recent macro-finance literature. Theorists have produced models with occasionally binding financial constraint. Once finance constraints are present (i.e., times of high financial stress), an amplification mechanism is triggered which increases the probability of observing a negative outcome (i.e., downside risk). We propose a way to test this prediction within our univariate and Bayesian framework, by incorporating a financial conditioning variable that shifts the transition probabilities between two states in the spirit of Diebold, Lee, and Weinbach (1994) and Filardo (1994). We show that the inclusion of a measure of financial conditions into the model helps to identify a regime characterized by downside risks and to predict when the economy may switch into it, thus providing evidence supporting recent macro-finance theories.

Related literature. As previously mentioned, there is a long tradition in macroeconomics in investigating the evolution of the distribution of the growth rate of macroeconomic time series data across various phases of the business cycle, i.e., expansion and recession phases. The first works come from Mitchell (1927), Keynes (1936), and Burns and Mitchell (1946), where they all provide purely descriptive statistical evidence on important first-moment asymmetries across phases. The econometric work by Hamilton (1989) represents a major step forward, by modeling and testing this idea. Since then, asymmetries in the first moment dynamics has been intensively studied within a modeling framework. Much of the existing empirical literature has also analysed the second moment of macroecononmic variables and thus has proposed symmetric measures of risk. For example, Baker, Bloom, and Davis (2016) proposes a measure of economic policy uncertainty using large databases of newpaper articles. Jurado, Ludvigson, and Ng (2015) and Clark, Carriero, and Massimiliano

(2017) estimate the time-varying volatility using a large number of macroeconomic variables. A major difference with all of these studies is that we depart from a symmetric distribution to emphasize the importance of asymmetric distribution.

More recently, there have been other studies that examine explicitely the time-varying tails of the distribution of GDP growth. Bloom, Guvenen, and Salgado (2016) analyse the behavior of skewness of macro- and micro-level variables over the business cycle. The authors compute the degree of skewness over time by using Kelly's measure. They report the dispersion in growth GDP during recessions is mostly driven by a decrease in the skewness, which in turn means a widening of the left tail of the distribution. Adrian, Boyarchenko, and Giannone (forthcoming) model empirically the conditional distribution of GDP growth as a function of economic and financial conditions using a two-step procedure that combines quantile regressions and skewed t-distributions. The authors report that recessions are associated with left-skewed distributions while, during expansions, the conditional distribution is closer to being symmetric. The present work should be regarded as largely complementary to these papers as the former uniquely focuses on simple statistical descriptive evidence based on suitable quantile based measures, as opposed to modern Bayesian tools we use here, and the two-step procedure of the latter is potentially slower and more costly relative to our Markov-switching framework. As said previously, 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. Our approach differs as we do not characterize the empirical effects of third-moment shocks on aggregate activity, but our empirical finding corroborates with the countercyclicity generated by their dynamic models.

From a methodological point of view, our paper is related to the vast literature on the use of Markov-switching models to macroeconomics. For example, univariate models include Hamilton (1989), Albert and Chib (1993), Diebold and Rudebusch (1996), Kim and Nelson (1999), and Morley and Piger (2012). Multivariate models include Sims and Zha (2006), Hubrich and Tetlow (2015), and Lhuissier (2017) for vector autoregressions (VAR), and Liu, Waggoner, and Zha (2011), Bianchi (2013), Lhuissier and Tripier (2016), Bianchi and Melosi (2017), and Lhuissier (2018) for DSGE models. Our work differ from previous contributions in the sense that they all allow the mean and/or shock variances to vary over time, and thus rule out, by construction, the possibility of skewed distributions.

The paper is organised as follows. Section II presents a brief overview of the skew-normal family of distributions. Section III outlines the Markov-switching model with skew-normal distributions, and explains how to estimate it. Section IV presents a MCMC method to carry out posterior inference. Section V applies our general framework to the growth rates of postwar U.S. real GDP, and presents the main results. Section VI conducts several exercises to assess the robustness of the results. Section VII tests the prediction that downside risk of GDP growth are associated with tighter financial conditions. Section VIII concludes.

II. The Skew-Normal Distribution: A preliminary

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. The following paragraphs provide the general framework of such distribution.

Let Y a random variable with the following density

$$p(Y|\xi, \sigma^2, \alpha) = \frac{2}{\sigma}\phi\left(\frac{Y-\xi}{\sigma}\right)\Phi\left(\alpha\frac{Y-\xi}{\sigma}\right), \tag{1}$$

where $\phi(.)$ and $\Phi(.)$ denote the standard normal density function and cumulative distribution function, respectively. We say that the random variable Y follows a univariate skew-normal distribution with location parameter ξ, scale parameter σ^2, and a skewness parameter α:

$$\text{skew-normal}(Y|\xi, \sigma^2, \alpha). \tag{2}$$

If the skewness parameter is equal to zero, then the density of Y is a normal distribution with mean ξ, and standard deviation σ.

The moments of the skew-normal distribution can be summarized as follows

$$\mathrm{E}[Y] = \xi + \sigma\delta\sqrt{\frac{2}{\pi}}, \qquad \mathrm{var}[Y] = \sigma^2\left(1 - \frac{2}{\pi}\delta^2\right), \tag{3}$$

where $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$ and $\delta \in (-1, 1)$.

As an illustration, Figure 1 displays skew-normal density functions when $\alpha = 0, -1, -4, -10$ in the left-hand panel, and $\alpha = 0, 1, 4, 10$ in the right-hand panel. For the remaining parameters, we set $\xi = 0$ and $\sigma^2 = 1$. As can be seen, the skewness parameter strongly alters the tails of the distribution. When α is negative, the distribution tends to be skewed to the left, while when it is positive, the distribution is skewed to the right.

An interesting characteristic of the skew-normal distribution is that it can be represented stochastically. In particular, the skew-normal distribution in (2) is equivalent to

$$Y = \xi + \delta Z + \sqrt{(1-\delta^2)}U, \tag{4}$$

where Z and U are random variables defined, respectively, as follows:

$$Z = \text{truncated-normal}(Z|0, \sigma^2)_{Z>0} \qquad \text{and} \qquad U = \text{normal}(U|0, \sigma^2), \tag{5}$$

with truncated-normal$(x|\mu, \Sigma)$ denotes the truncated-normal distribution of x with mean μ, variance Σ, and truncation below zero, and normal$(x|\mu, \Sigma)$ denotes the normal distribution of x with mean μ and variance Σ. Say it differently, the skew-normal distribution may be seen as the combination of a normal random variable and a truncated standard normal variable.

In next sections, we will show that this elegant and stochastic representation is crucial in order to obtain our Gibbs-sampling procedure.

III. Autoregressive models with Markov Skewness Shifts

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

$$y_t = c + \phi_1(y_{t-1} - c) + \ldots + \phi_\tau(y_{t-\tau} - c) + \epsilon_t, \qquad t = 1, \ldots, T \tag{6}$$

where ϕ's contain the coefficients at the lag τ; c is a constant; and T is the sample size. We assume that ϵ_t follows a skew-normal distribution as

$$\text{skew-normal}(\epsilon_t|0, \sigma^2, \alpha(s_t)), \tag{7}$$

where σ and $\alpha(s_t)$ represents the scale and the shape parameters, respectively. For $1 \leq i, j \leq H$, the discrete and unobserved variable s_t that governs the shape parameter is an exogenous first order Markov process with the transition matrix Q

$$Q = \begin{bmatrix} q_{1,1} & \cdots & q_{1,j} \\ \vdots & \ddots & \vdots \\ q_{i,1} & \cdots & q_{i,j} \end{bmatrix}, \tag{8}$$

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

Define $Y_t = [y_1, \ldots, y_t]$, and $x_t = [y_{t-1} - c, \ldots, y_{t-\tau} - c, 1]'$ for $t \geq 1$, and let $\theta = (\phi, \sigma^2, \alpha(k), Q)$, where $\phi = [\phi_1, \ldots, \phi_\tau, c]$ and $k = \{1, \ldots, H\}$. Then, the conditional likelihood at time t is given by

$$p(y_t|Y_{t-1}, s_t, \theta), \tag{9}$$

which is generated by

$$p(y_t|Y_{t-1}, s_t, \theta) = \frac{2}{\sigma}\phi\left(\frac{y_t - \phi x_t}{\sigma}\right)\Phi\left(\alpha(s_t)\frac{y_t - \phi x_t}{\sigma}\right). \tag{10}$$

Given (9), it follows that the overall likelihood of Y_T is

$$p(Y_T|\theta) = \prod_{t=1}^{T}\left[\sum_{s_t \in h} p(y_t|Y_{t-1}, s_t, \theta)\Pr(s_t, \theta)\right]. \tag{11}$$

The object inside the brackets of the likelihood in (11) 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 (See Appendix A). Interestingly, the inclusion of the additional shape parameter does not require to modify the original filter.

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

$$p(\theta|Y_T) \propto p(Y_T|\theta)p(\theta), \tag{12}$$

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

For computational reasons, we employ a logarithm transformation in equation (12) to obtain the log-posterior function as follows

$$\log\{p(\theta|Y_T)\} \propto \log\{p(Y_T|\theta)\} + \log\{p(\theta)\}, \tag{13}$$

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

$$\log\{p(y_t|Y_{t-1}, s_t, \theta)\} = \text{constant} - \log\{\sigma\} - \frac{(y_t - \phi x_t)^2}{2\sigma^2} + \log\left\{\Phi\left(\alpha(s_t)\frac{y_t - \phi x_t}{\sigma}\right)\right\}. \tag{14}$$

The strategy to find the posterior mode of (13) 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 us to correctly cover the parameter space and avoid getting stuck in a "local" peak. Note, however, that we 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), in which the authors break the parameters into several subblocks of parameters and 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 in (6) remains relatively small and allows us employ a standard technique.

IV. A Gibbs sampler

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). We work differs from this literature along several dimensions. First, we 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, we introduce an autoregressive process of finite order, as naturally modelled in the macroeconomics literature. Third, our MCMC algorithm 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. Fourth, our method is able to deal with independently switches between the mean, the scale and the shape parameters over time, while Frühwirth-Schnatter and Pyne (2010) allow only switches in a synchronized manner. Overall, Our MCMC approach can be seen as an extension of Albert and Chib (1993) to Markov skewness shifts.[2]

A MCMC simulation method is employed to approximate the joint posterior density $p(\theta, Z_T, S_T|Y_T)$, where $S_t = [s_1, \ldots, s_t]$, and $Z_t = [z_1, \ldots, z_t]$ for $t \geq 1$. Here, a key to Bayesian estimation of a Markov-switching skewed model is to apply a stochastic representation of equations (6) and (7) as follows

$$y_t = \phi x_t + \delta(s_t) z_t + \sqrt{1 - \delta(s_t)^2}\nu_t, \qquad t = 1, \ldots, T \tag{15}$$

where z_t and ν_t are random variables at time t defined, respectively, as follows:

$$\text{truncated-normal}(z_t|0, \sigma^2)_{z_t>0} \qquad \text{and} \qquad \text{normal}(\nu_t|0, \sigma^2). \tag{16}$$

Because we consider a Bayesian approach to model (15) and (16), we now explicit our priors. The prior on the set of parameters θ is given by:

$$\phi = \text{normal}(\phi|\bar{b}, \bar{B}), \tag{17}$$

$$\sigma = \text{inv-gamma}(\sigma|\bar{\alpha}, \bar{\beta}), \tag{18}$$

$$q_k = \text{dirichlet}(q_k|\bar{\alpha}_{1k}, \ldots, \bar{\alpha}_{hk}), \tag{19}$$

$$\alpha(k) = \text{normal}(\alpha(k)|\alpha_0, \psi_0), \tag{20}$$

where $\bar{b}, \bar{B}, \bar{\alpha}, \bar{\beta}$ and $\bar{\alpha}_{1k}, \ldots, \bar{\alpha}_{hk}$ are the hyperparameters; and dirichlet$(q_k|\alpha_1, ..., \alpha_h)$ is the Dirichlet distribution of q_k as follows:

$$\frac{1}{B(\alpha)} \prod_{i=1}^{h} {q_i}^{\alpha_i - 1} \tag{21}$$

with $B(\alpha) = \frac{\prod_{i=1}^{h} \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^{h} \alpha_i\right)}$, where Γ denotes the standard gamma function. As can be seen, we directly specify informative priors for the shape parameter $\alpha(k)$ rather than for $\delta(k)$, the transformed shape parameters.[3]

[2]Albert and Chib (1993) develop a Gibbs sampling for AR time series subject to Markov mean and variance shifts.

[3]When specifying priors for $\delta(k)$, instead of $\alpha(k)$, there is no closed form for the posterior distribution of $\delta(k)$, 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 (15) and (16) lead us to exploit the idea of Gibbs-sampling. Let $\theta_{\neq x}$ contain the model's parameters, except for x. The MCMC sampling scheme at the (i)st iteration, for $i = 1, \ldots, N_1 + N_2$, consists of sampling from the following conditional posterior distributions

(1) $p\left(S_T^{(i)} | Y_T, \theta^{(i-1)}\right)$;
(2) $p\left(Q^{(i)} | S_T^{(i)}\right)$;
(3) $p\left(Z_T^{(i)} | Y_T, S_T^{(i)}, \theta^{(i-1)}\right)$;
(4) $p\left(\phi^{(i)} | Y_T, S_T^{(i)}, Z_T^{(i)}, \theta_{\neq\phi}^{(i-1)}\right)$;
(5) $p\left(\sigma^{(i)} | Y_T, S_T^{(i)}, Z_T^{(i)}, \phi^{(i)}, \delta^{(i-1)}\right)$;
(6) $p\left(\alpha^{(i)} | Y_T, S_T^{(i)}, Z_T^{(i)}, \theta_{\neq\alpha}^{(i)}\right)$.

A few items deserve discussion. First, simulation from the conditional posterior density $p\left(S_T^{(i)} | Y_T, \theta^{(i-1)}\right)$, given Z_T and θ, is standard and in closed form. Second, simulation from the conditional posterior density $p\left(Q^{(i)} | S_T^{(i)}\right)$ is independent of the time series Y_T, the random variable Z_T and the model's other parameters. Third, simulation from the conditional posterior density $p\left(Z_T^{(i)} | Y_T, S_T^{(i)}, \theta^{(i-1)}\right)$, given Y_t, Z_t and θ, is available in closed form due to the stochastic representation of the Markov-switching model through normal and truncated-normal variables. Fourth, simulations from the conditional posterior densities $p\left(\phi^{(i)} | Y_T, S_T^{(i)}, Z_T^{(i)}, \theta_{\neq\phi}^{(i-1)}\right)$ and $p\left(\sigma^{(i)} | Y_T, S_T^{(i)}, Z_T^{(i)}, \phi^{(i)}, \delta^{(i-1)}\right)$ reduces to Bayesian inference for Markov-switching models with known allocations, S_T. Finally, simulation from the conditional posterior density $p\left(\alpha^{(i)} | Y_T, S_T^{(i)}, Z_T^{(i)}, \theta_{\neq\alpha}^{(i)}\right)$ is in closed form, and follows an unified skew-normal distribution introduced by Arellano-Valle and Azzalini (2006).

This sampler begins with setting parameters at the peak of the posterior density function. We 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. With respect to Albert and Chib (1993), our Gibbs-sampling procedure involves two more blocks, namely the conditional posterior distribution of Z_T, given the parameters and the states, and the conditional posterior distribution of $\alpha(k)$, given Z_T, S_T and the remaining parameters.

The researcher can use our companion computer program[4], written in C++, to estimate and simulate an AR model with Markov skewness shifts. The user just needs to provide an

[4]The software is available at http://stephanelhuissier.eu/assets/skewcodes.zip. In Appendix B, we provide a concrete example of how to use the interface with our C++ computer code. The software program was written in modern C++11 and mainly uses the GNU Scientific Library (GSL-2.5) and the Eigen library (3.3.5). Both libraries are open source.

input file in which he/she must mention each specification (such as the number of lags, prior settings, the number of draws, the number of burn-in, etc...) of the AR model. Due to its simplicity and efficiency, we believe that our companion computer code is relevant for anyone interested in inference of such models.

The subsections that follow provide the computational details for each conditional posterior distribution.

IV.1. **Conditional posterior densities,** $p\left(S_T^{(i)}|Y_T,\theta^{(i-1)}\right)$**.** For $t=1,2,...,T$, we can generate $S_T^{(i)}$ using the Carter and Kohn (1994)'s multi-move Gibbs-sampling as following

$$p(S_T^{(i)}|Y_T,\theta^{(i)}) = p(s_T^{(i)}|Y_T,\theta^{(i)})\prod_{t=1}^{T-1} p(s_t^{(i)}|S_{t+1}^{(i)},Y_T,\theta^{(i)}). \quad (22)$$

Drawing $S_T^{(i)}$ from the full conditional distribution based on this equation is standard. We 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 $s_{T-1}, s_{T-2}, \ldots, 1$ according to

$$p(s_t|Y_T,\theta) = \sum_{s_{t+1}} p(s_t|Y_T,\theta,s_{t+1})p(s_{t+1}|Y_T,\theta), \quad (23)$$

where

$$p(s_t|Y_T,\theta,s_{t+1}) = \frac{\Pr\left[s_{t+1}|s_t\right]p(s_t|Y_T,\theta)}{p(s_{t+1}|Y_t,\theta)}. \quad (24)$$

Appendix A provides the details for derivation of the Hamilton (1989) filter.

IV.2. **Conditional posterior densities,** $p\left(Q^{(i)}|S_T^{(i)}\right)$**.** The conditional posterior distribution of $Q^{(i)}$ is as follows:

$$p(q_k^{(i)}|S_T) = \text{dirichlet}(q_k^{(i)}|\bar{\alpha}_{1k}+n_{1k},...,\alpha_{\bar{H}k}n_{Hk}) \quad (25)$$

where $q_k^{(i)}$ is the kth column of $Q^{(i)}$, n_{ij} is the total number of transitions from state j to state i over the entire sample.

Drawing $Q^{(i)}$ from the above full conditional distribution is also standard.

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

For $t=1,2,...,T$, we generate $Z_T^{(i)}$ according to

$$p\left(Z_T^{(i)}|Y_T,S_T^{(i)},\theta^{(i-1)}\right) = \prod_{t=1}^{T} p\left(z_t^{(i)}|Y_t,S_t^{(i)},\theta^{(i-1)}\right), \quad (26)$$

where

$$p\left(z_t^{(i)}|Y_t, S_t^{(i)}, \theta^{(i-1)}\right) = \text{truncated-normal}(z_t^{(i)}|\delta(s_t^{(i)})(y_t - \phi^{(i)} x_t), \sigma^2(1-\delta(s_t)^{(i)^2}))_{z_t^{(i)}>0}. \quad (27)$$

IV.4. **Conditional posterior densities,** $p\left(\phi^{(i)}|Y_T, S_T^{(i)}, Z_T^{(i)}, \theta_{\neq\phi}^{(i-1)}\right)$**.** If we let $y_t^* = \frac{y_t - \delta(s_t) z_t}{\sqrt{1-\delta(s_t)^2}}$, and $x_t^* = \frac{x_t}{\sqrt{1-\delta(s_t)^2}}$, we obtain an homoskedastic model as follows

$$y_t^* = \phi x_t^* + \nu_t, \quad (28)$$

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

$$p\left(\phi^{(i)}|Y_T, S_T^{(i)}, Z_T^{(i)}, \theta_{\neq\phi}^{(i-1)}\right) = \text{normal}\left(m_\phi^{(i)}, M_\phi^{(i)}\right), \quad (29)$$

where

$$m_\phi^{(i)} = \left(\bar{B}^{-1} + X'X\right)^{-1}\left(\bar{B}^{-1}\bar{b} + X'y_t^*\right), \quad (30)$$

$$M_\phi^{(i)} = \left(\bar{B}^{-1} + X'X\right)^{-1}, \quad (31)$$

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

IV.5. **Conditional posterior densities,** $p\left(\sigma^{(i)}|Y_T, S_T^{(i)}, Z_T^{(i)}, \phi^{(i)}, \delta^{(i-1)}\right)$**.** Given Y_t, S_T, Z_T, θ, and S_T, the scale parameter σ can be drawn using the following inverse-gamma distribution

$$p\left(\sigma^{(i)}|Y_T, S_T^{(i)}, Z_T^{(i)}, \phi^{(i)}, \delta^{(i-1)}\right) = \text{inv-gamma}(\bar{\alpha} + ssr, \bar{\beta} + T), \quad (32)$$

where ssr is the sum of squared residual defined as

$$ssr = \sum_{t=1}^{T}\left(\frac{y_t - \mu_t^{(i)} - \delta(s_t)^{(i)} z_t^{(i)}}{\sqrt{1-\delta(s_t)^{(i)^2}}}\right)^2, \quad (33)$$

where $\bar{\alpha}$ and $\bar{\beta}$ are the shape hyperparameters implied by the choice for the prior mean and variance.

IV.6. **Conditional posterior densities,** $p\left(\alpha^{(i)}(k)|Y_T, S_T^{(i)}, \theta_{\neq\alpha}^{(i)}\right)$**.** Let $\bar{y}_t = \frac{y_t - \phi x_t}{\sigma}$ and $\overline{Y}_T = [\bar{y}_1, \ldots, \bar{y}_T]'$. Consider the following derivation for the full conditional distribution of $\alpha(k)$, given Y_T, $S_T^{(i)}$, and $\theta_{\neq\alpha}^{(i)}$:

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

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

$$\phi_d\left(z-\xi; \omega\Omega\omega\right)\frac{\Phi_m\left(\gamma + \Delta\Omega^{-1}\omega^{-1}(z-\xi); \Gamma - \Delta\Omega^{-1}\Delta'\right)}{\Phi_m(\gamma;\Gamma)^{-1}}, \tag{34}$$

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 $\zeta_t = \psi_0\bar{y}_t^2(\psi_0^2\bar{y}_t^2 + 1)^{-1/2}$; $\Gamma_1 = I - \mathrm{diag}(\Delta_1)^2 + \Delta_1\Delta_1^2$; and where $\mathrm{diag}(V)$ is a diagonal matrix, the elements of which coincide with those of vector V.[5]

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

$$U_0 = \text{normal}\left(U_0|0, \overline{\Psi}_\Delta\right), \quad \text{and} \quad U_{1,-\gamma} = \text{truncated-normal}(U_1|0,\Gamma)_{-\gamma}. \tag{35}$$

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

$$\xi + \omega\left(U_0 + \Delta\Gamma^{-1}U_{1,-\gamma}\right). \tag{36}$$

Once we obtain $\alpha(k)$, we can directly transform it to recover $\delta(k) = \frac{\alpha(k)}{\sqrt{(1+\alpha(k)^2)}}$. Due to the label-switching problem, we normalize the labels of regimes to obtain accurate posterior distributions as follows $\alpha(1) < \ldots < \alpha(H)$. To achieve this constraint, we adopt rejection sampling.

[5]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.

V. GDP Growth Rates

In this section, we investigate the time-varying skewness of the U.S. real GDP growth rates constructed from the seasonnally adjusted quarterly real GDP series for the period 1952.Q1—2016.Q1. The GDP growth rates (GDPC1) come from the St. Louis Federal Reserve's database — FRED (Federal Reserve Economic Data).[6] The time series is presented in Figure 2. As a starting point, we consider an autoregressive Markov-switching model of order 1. Our AR(1) Markov-switching model is as follows

$$y_t = c + \phi_1 y_{t-1} + \epsilon_t, \qquad t = 1, \ldots, T \tag{37}$$

where T is the sample size; c denotes the constant term; y_t is the log difference of real GDP; y_{t-1} is the log difference of lagged real GDP; ϕ_1 is the corresponding regression parameter and ϵ_t follows a skew-normal distribution as

$$p(\epsilon_t | Y_{t-1}, Z_t, S_t, \theta) = \text{skew-normal}(\epsilon_t | 0, \sigma^2, \alpha(s_t)), \tag{38}$$

with σ is the scale parameter, and s_t is the exogenous first-order two-state Markov process defined in (8), where $H = 2$. For the reasons explained above, we transform equation (37) into a stochastic representation as follows

$$y_t = c + \phi_1 y_{t-1} + \delta(s_t) z_t + \sqrt{1 - \delta(s_t)^2} \nu_t, \qquad t = 1, \ldots, T, \tag{39}$$

where $\delta(k) = \frac{\alpha(k)}{\sqrt{1+\alpha(k)^2}}$ for $1 \leq k \leq H$, and z_t and ν_t are random variables at time t defined in (16).

We estimate and simulate our AR(1) model with Markov skewness shifts for the GDP growth rates using our maximization and MCMC procedures developed in Sections III and IV, respectively.

The priors are defined in Table 1, which reports the specific distribution, the mean and the standard deviation for each parameter. A few of them deserve further discussion. First, for c and ϕ_1 we choose a normal prior with the mean 0.00 and the standard deviation 5.00. The prior for the scale parameter (σ) follows an inverse-gamma distribution, with the mean 1.00 and the standar deviation 1.00. These two priors are rather dispersed and cover a large parameter space. The prior for the shape parameters ($\alpha(k), k = 1, 2$) has a normal density with the mean 0.00 and the standard deviation 3.00. It may be worth noting that we impose the exact same prior across regimes, so that differences between shape parameters result from data rather than priors. Finally, the prior duration of each regime is about five quarters, meaning that the average probability of staying in the same regime is equal to 0.80 and a standard deviation equal to 0.10.

[6]The real GDP series is available here: https://fred.stlouisfed.org/series/GDPC1

The results shown in this paper are based on $11,000$ draws with our Gibbs-sampling procedure developed in Section IV. We discard the first $1,000$ draws as burn-in, and keep every 10-th draw in order to achieve an approximately independent sample. On the right-hand side of Table 1, we report the posterior mode, mean, and median with the 90 percent probability interval for each parameter of the Markov-switching AR(1) model. Figure 4 displays marginal posterior density estimates for parameters of the model using normal kernel density estimates.

The first finding that is evident is the remarkable difference in the estimated shape parameters across the two states. The first state gives a value close to zero to the shape parameter (0.1496 at the mode), while its value in the second state is positive (3.5034 at the mode). The probability intervals for $\alpha(1)$ lies within both the negative and positive regions, suggesting that there is no particular asymmetry in the first regime. By contrast, the probability interval for $\alpha(2)$ is tightly concentrated within the positive region but also skewed to the right, which reinforces our estimates and reveals a regime marked by positive skewness. For all of these reasons, we label "Regime 1" as the non-skewed distribution regime, and "Regime 2" as the positively (or right) skewed distribution regime.

Regarding the posterior probabilities (q_{11} and q_{22}) of the Markov-switching process, it is apparent that the persistence of staying in each state is relatively high. The 90% probability intervals for q_{11} are 0.8440 and 0.9146, and those for q_{22} are 0.8440 and 0.9560, indicating that the non-skewed distribution regime is much less persistent (an average duration of 6 quarters at the mode) than the right skewed distribution regime (an average duration of 12 quarters at the mode). Once again, posterior means and medians are concentrated in tight ranges, reinforcing our estimated parameters. In addition to that, we also report how robust are the differences in the estimates between the two regimes by looking at marginal posterior density estimate. When looking at the bottom right corner of Figure 4 or the last row of Table 1, the distribution of $\alpha(2) - \alpha(1)$ is entirely displayed in the positive region, meaning that the differences between regimes are apparent. Results like this clearly imply that macroeconomic models should go beyond both linearity and gaussianity to better understand business cycle fluctuations.

Figure 8 reports the (filtered) probabilities — evaluated at the mode — of being in Regime 1 over time.[7] One can see from the figure that the U.S. economy has been characterized by switches between the two regimes regimes over time. In particular, the non-skewed regime coincides remarkably well with the recessions dates declared by the NBER's Business Cycle Dating Committee. Say it differently, the distribution of the quarterly growth rate of GDP

[7]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.

tends to be symmetric in recession phases, while the distribution is skewed to the right in expansion phases.

VI. Model comparison and alternative approaches

In Section VI.1, we compare our benchmark model to other competing models via a Bayesian model selection criteria. In Section VI.2, we employ alternative approaches to assess the robustness of our finding.

VI.1. **Model comparison.** We employ the Watanabe-Akaike Information Criterion (WAIC), introduced by Watanabe (2010), for purposes of model comparison. WAIC evaluates the predictive accuracy for a fitted model by computing the log pointwise predictive density corrected from the effective number of parameters to adjust for overfitting. WAIC offers two main advantages. First, it is fully Bayesian in that it is based on the usual posterior simulations of the parameters. Second, it is invariant to parametrization.

We model, estimate, and compare several competing models. First, we consider autoregressive models with Markov skewness shifts of different order (0, 2, 3, and 4).[8] Second, we consider the constant version of our benchmark model, i.e., a constant-parameters autoregressive model with skew-normal errors ("Constant-skew normal AR(1)"). Third, we consider a constant-parameters autoregressive model with normal errors ("Constant normal AR(1)"). Fourth, we consider the Hamilton (1989) autoregressive model, in which the mean of the GDP growth can change over time according to a two-states Markov-switching process, thus generating dynamics qualitatively different in expansionary and contractionary periods ("MS Hamilton AR(1)").[9] In this section, we want to know whether AR(1) model with Markov skewness shifts represents the best description for the U.S. real GDP growth.

[8]In Appendix C and D, we report the filtered probabilities and the posterior moments for each AR model with Markov skewness shifts.

[9]More specifically, the Hamilton (1989) model allows the mean of the GDP growth rate to be evolving as follows:

$$
\begin{aligned}
y_t - \mu(s_t) &= \phi_1(y_{t-1} - \mu(s_t)) + \varepsilon_t, \\
p_{i,j} &= \Pr(s_t = i | s_{t-1} = j), \qquad i, j = 1, 2,
\end{aligned}
$$

where the distribution for ε_t is as follows: $p(\varepsilon_t) = \text{normal}(\varepsilon_t | 0, \sigma_{\varepsilon,t})$. The priors for $\mu(k)$ is chosen so be asymmetric across regimes, so that the dynamics of recessions are qualitatively distinct from those of expansions. More specifically, the prior for $\mu(1)$ follows a normal distribution with mean -0.50 and standard deviation 5.00, and the prior for $\mu(2)$ follows a normal distribution with mean 1.00 and standard deviation 5.00. Note, however, that our priors are loose enough to let the data dominate the posterior distribution. We estimate the process by employing the Kim and Nelson (1999)'s Gibbs sampling procedure to alternately sample from the conditional distributions.

Table 2 displays the value and the standard deviation of WAIC for each model, as well as their differences relative to our benchmark AR(1) model with Markov skewness shifts. When comparing two estimated models, one must also look at not only the estimated difference in their expected predictive accuracy (column 4), but also the standard error of this difference (column 5), which gives a good summary of uncertainty.

Regarding the AR models, AR(0) model might be tempting as the value of its WAIC is the lowest; the log difference in WAIC between the AR(0) model and our benchmark AR(1) model is 2.01. However, as said previously, it is important to look at the standard error of their difference (i.e., last column of the table), which is twice as large as their difference, meaning that the AR(0) model does not outperform our benchmark model, but is also a good candidate to describe the data. Overall, all of AR models are best-fit models. As shown in Appendix C and D, the results from these models are quite similar.

Regarding constant competing models, the levels of their WAIC are much higher than the benchmark's one. Indeed, their estimated WAIC differences range from -30.59 to -29.04, with standard errors equal to about 6. Although the Hamilton (1989)'s model offers a lower estimated WAIC than the constants' one, it fails to outperform our benchmark model. Note, however, that this does not mean our benchmark model is at odds with the finding of Hamilton (1989), i.e., the mean of GDP growth is qualitatively different in recessions than in expansions, as the shape parameter affects also the mean of skew-normal distributions, as shown in equation (3).

Overall, we conclude that the AR(1) model with Markov skewness shifts is a convenient description of GDP growth.

VI.2. **Alternative approaches.** The analytical framework of this paper may be potentially problematic as it may not identify the correct source of time variation in the distribution of GDP growth rates over the business cycle. Indeed, suppose that the differences between recessions and expansions reflect in reality changes only in the location parameter of the AR process. For example, Hamilton (1989) finds that, by estimating an AR(4) switching model to U.S. real GDP series, abrupt changes in the mean of the GDP growth rate over time due to the time-varying location parameter.

Suppose further that the model kept the location parameter invariant and allowed only the shape parameter to drift over time, as in our modeling framework. Since the location parameter is held constant, the estimated model would fail to detect any drift in the level. However, because the time-varying skewness model is a serious candidate for the explanation of changes in the distribution, the skewness is likely to change in order to compensate for the absence of variations in the level.

Another reason why the skewness switching model might be incorrect is that it does not properly into account hetereskedasticity, which can be captured by time-variation in the scale parameter. For example, Sims (2001) and Lhuissier and Zabelina (2015) have shown the importance of capturing changes in volatility before allowing changes in remaining parameters in order to avoid misleading results. Among others, Kim and Nelson (1999), McConnell and Perez-Quiros (2000), and Stock and Watson (2003) have provided strong evidence of a structural break in the postwar U.S. real GDP growth rates in the early 1980s. As a consequence, we think it is reasonable to identify the structural break in volatility when making allowance for skewness shifts.

The objective of this section is to provide evidence that our main results remain unchanged when allowing changes in the level of GDP growth rates or changes in the volatility over time. In section VI.2.1, we estimate a Markov-switching model in which both the mean and the skewness of real GDP growth rates can switch over time. In section VI.2.2, we estimate our skewness switching model with a deterministic single break in the scale parameter in 1984.Q4.

VI.2.1. *Changes in the location parameter.* We extend our benchmark model by allowing both the location and shape parameters to switch over time according to a Markov-switching process. To do so, we replace equation (37) as follows

$$y_t = c(s_t) + \phi_1 y_{t-1} + \epsilon_t, \qquad t = 1, \dots, T,$$

where ϵ_t follows the same distribution as equation (7). Put it differently, the times of location changes are stochastically dependent of the times of shape changes. Once again, we estimate the parameters of the AR process with our Bayesian methods. The MCMC procedure is slightly modified as here we also allow the shape parameter to vary over time. We do not report the new procedure as the modification is straightforward.

Regarding the priors, we choose exactly the same as those previously used, except for $c(s_t)$, for which we choose asymmetric priors. For the first regime, we impose a normal prior with the mean 0.00 and the standard deviation 5.00, while for the second regime, we choose a mean equal to 1.00 and a standard deviation to 5.00. This is meant to reflect that the mean of U.S. real GDP growth rates is qualitatively different in recession than in expansion phases. However, it is important to notice that the priors remain very loose, implying that the data, and therefore the likelihood, dominate the posterior distribution.

Table 3 reports priors, modes, means, medians and 90% probability intervals for the AR(1) model with Markov location and shape shifts. Results reveal that there are still important differences across regimes, where the estimate for $\alpha(1)$ is equal to zero at the mode with the probability interval $[-5.1699; 1.0876]$ in Regime 1, while its estimate is about 4.2200 with probability interval $[1.7861; 7.1306]$ in Regime 2. Interestingly, it turns out that there are not any major differences in the location parameter across regimes. Indeed, the 90% error bands for each $c(k)$ lie within the negative and positive regions, indicating that both parameters are similar across regimes.

Finally, although not reported[10], the estimated (filtered and smoothed) probabilities from this model are similar to those from the benchmark model, namely regime switching occurs between periods of recessions and expansions. Results like this reveal that changes in the distribution in the growth rates of postwar U.S. real GDP over the business cycle result from important changes in risk tails.

VI.2.2. *Structural Break.* We now estimate the model across two subsamples: 1952.Q1-1984.Q4 and 1985.Q1-2018.Q2. By doing so, we are able to take into account the structural break in 1984.Q4, as noted in Kim and Nelson (1999).

Estimation results are reported in Tables 4 (throughout the pre-Great Moderation) and 5 (throughout the post-Great Moderation). As it can be seen, the fact of taking into account heteroskedasticity does not really affect the main conclusions. There are still important differences in estimates of the skewness parameters across the two regimes, where its mode is equal to $\alpha(1,1) = -0.3263$ and $\alpha(1,2) = -0.0029$ in Regime 1, and $\alpha(2,1) = 1.5721$ and $\alpha(2,2) = 2.5912$ in Regime 2. The 90% probability intervals for $\alpha(2,1) - \alpha(1,1)$ and $\alpha(2,2) - \alpha(1,2)$, which are equal to $[0.2790; 4.9478]$ and $[1.6313; 8.3641]$ respectively, still indicate that major differences between regimes are apparent. Note also that these estimates are very close between the pre- and post- Great Moderation periods. Finally, the estimates for $\sigma(1)$ and $\sigma(2)$ at the mode are equal to 1.09 and 0.62, with tight probability intervals. The higher degree of volatility in the pre-1984 period corroborates , for example, with Kim and Nelson (1999).

Summarising, there is strong evidence of changes in the tails of distribution of GDP growth over the business cycle and this is clearly a robust phenomenon.

VII. GDP skewness and financial conditions

In the past few years, theories in macro-finance literature has made much progress in understanding the relation between financial conditions and the distribution of GDP growth. Recent research (e.g., He and Krishnamurthy, 2012, 2013; Adrian and Boyarchenko, 2012; Brunnermeier and Sannikov, 2014; Maggiori, 2017) has produced models in which the intermediary sector is occasionally constraint, thus generating non-linear amplification. Intuitively, while in normal times the good financial conditions of financial intermediaries can

[10]All results of this section are available upon available.

absorb losses induced by a negative shock, in times of crisis the financial sectors fragility creates grave and long lasting problems with firms financing conditions tightened, leading to substantial cuts in spending, employment and production. Say it differently, disturbances to the financial intermediation sector potentially cause downside risks of GDP growth (i.e., greater negative skewness). With a few exceptions[11], the empirical evidence is largely absent from the literature.

In this section, we test the prediction that downside risks of GDP are associated with tighter financial conditions by extending our Markov-switching skewed AR model in Section III to incorporate a financial conditioning variable that shifts the transition probabilities between two states in the spirit of Diebold, Lee, and Weinbach (1994), Filardo (1994), and Filardo and Gordon (1998).

VII.1. **Extension of the model to allow time-varying transition probabilities.** To construct our formal test, we replace the fixed transition probabilities (FTP) in equation (8) by a time-varying transition probabilities (TVTP) as follows:

$$Q = \begin{bmatrix} q_{1,1}(z_t) & 1 - q_{2,2}(z_t) \\ 1 - q_{1,1}(z_t) & q_{2,2}(z_t) \end{bmatrix}, \tag{40}$$

where z_t is a lagged financial conditioning variable. Specifically, we use the National Financial Conditions Index (NFCI) computed by the Federal Reserve Bank of Chicago.[12]

We use a univariate probit model to measure the evolution of the unobservable regime s_t:

$$\Pr[s_t = 1] = \Pr[s_t^* \geq 0], \tag{41}$$

where s_t^* is a latent variable defined as:

$$s_t^* = \gamma_0 + \gamma_1 s_{t-1} + \beta_0(1 - s_{t-1})z_t + \beta_1 s_{t-1} z_t + u_t, \tag{42}$$

where u_t follows a standard normal distribution:

$$p(u_t) = \text{normal}(u_t|0, 1). \tag{43}$$

The parameter β_0 and β_1 determine how our financial conditioning variable affects the transition probabilities of regimes. Thus, the transition probabilities $q_{1,1}(z_t)$ and $q_{2,2}(z_t)$ are given

[11]Recently, Adrian, Boyarchenko, and Giannone (forthcoming) find evidence that downside risks to U.S. GDP growth are predicted by financial conditions. Adrian, Grinberg, Liang, and Malik (2018) examine the distribution of expected GDP growth for 11 advanced economies and find similar results.

[12]See https://www.chicagofed.org/publications/nfci/index

by:

$$\begin{aligned} q_{1,1}(z_t) &= \Pr(s_t = 0 | s_{t-1} = 0, z_{t-1}] \\ &= \Pr[s_t^* < 0 | s_{t-1} = 0, z_{t-1}] \\ &= \Pr[u_t < -\gamma_0 - \beta_0 z_{t-1}] \\ &= \Phi(-\gamma_0 - \beta_0 z_{t-1}). \end{aligned} \tag{44}$$

$$\begin{aligned} q_{2,2}(z_t) &= \Pr[s_t = 1 | s_{t-1} = 1, z_{t-1}] \\ &= \Pr[s_t^* \geq 0 | s_{t-1} = 1, z_{t-1}] \\ &= \Pr[u_t \geq -\gamma_0 - \gamma_1 - \beta_1 z_{t-1}] \\ &= 1 - \Phi(-\gamma_0 - \gamma_1 - \beta_1 z_{t-1}). \end{aligned} \tag{45}$$

with $\Phi(.)$ represents the cumulative distribution function for u_t. When $\beta_0 = 0$ and $\beta_1 = 0$, we have fixed transition probabilities. 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}, and on financial conditions z_t. In other words, variations in z_t will affect the expected duration of how long a regime will last.

To let the data "decide" how important financial conditions are, we choose very loose priors as follows

$$p(\gamma_i) = \text{normal}(\gamma_i | 0, 2^2), \qquad p(\beta_i) = \text{normal}(\beta_i | 0, 2^2), \tag{46}$$

where $i = \{0, 1\}$. Here, γ_i and β_i follows a normal distribution with mean 0 and standard deviation equal to 2. As centered around zero, each of these prior distributions include the possibility that the financial conditioning variable z_t is excluded from the model, and thus that financial conditions do not affect transition probabilities.

VII.2. **Implementation of Gibbs-sampling.** The Gibbs-sampling procedure discussed in Section IV needs to be modified as the transition probabilities are now time-varying. Specifically, the procedure for generating S_T in Section IV.1 and Q in Section IV.2 require some modifications.

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. Conditional on S_T, γ_i, and β_i, one can generate Z_T based on equation (42) by generating u_t from an appropriate truncated standard normal distribution.

Conditional on S_T, S_T^*, and Z_t, one can easily generate draws for γ_i and β_i using equation (42), which represents a simple linear regression model. Suppose that $\theta \in (\gamma_0, \gamma_1, \beta_0, \beta_1)$, $W = [\mathbb{1} \quad Z_T \quad S_{T-1}]$, and $S_T^* = [S_1^*, \ldots, S_T^*]'$. The conditional posterior distribution is then

$$p(\theta|S_T, S_T^*, Z_T) = \text{normal}(m_\theta, M_\theta), \tag{47}$$

where $M_\theta = (\bar{B}_\theta^{-1} + W'W)^{-1}$, and $m_\theta = M_\theta(\bar{B}_\theta^{-1}\bar{b}_\theta + W'S_T^*)$. The closed form of the conditional posterior allows us to easily generate values for θ.

Once values for γ_i and β_i are generated, one can easily compute the transition probabilities $q_{1,1}(z_t)$ and $q_{2,2}(z_t)$ using the normal cumulative density function in equations (44) and (45), respectively.

Generating the other conditional distributions of the system is exactly the same as of fixed transition probabilities, as described in Section IV.

VII.3. **Empirical Results.** We estimate our augmented model with time-varying transition probabilities using data sample from 1970.Q2 to 2018.Q2. Because the NFCI is only available from 1970, we do not include the preceding period in our estimation. This augmented model is compared to the corresponding fixed transition probabilities specification, as discussed in preceding sections.

Figure 5 reports the (filtered) probabilities of being in Regime 1 over time produced by the TVTP (shown in the right panel) and FTP (shown in the left panel) models. Interestingly, the times of changes are different. First, the NBER recession in the early 2000 is only captured by the FTP model. Second, the expected durations tend to be shorter in the TVTP model. Third, the TVTP probabilities are near zero during the post-Great Recession, while they can exceeded 50 percent in the FTP model. All in all, the TVTP model appears to identify a different regime, via its leading indicator, from the one reported by the FTP model.

To better understand the differences between both models, Tables 6 and 7 report the posterior moments for both models. The most striking observation is the remarkable difference in the estimates for α in Regime 1 across TVTP and FTP models. Indeed, the TVTP model reports that the estimates for $\alpha(1)$ is -1.8931 with the probability interval $[-5.4079; -0.0200]$, while the FTP model indicates a more moderate negative skewness with -0.1093. Note also that the posterior distribution for $\alpha(1)$ in the FTP model tends to suggest that the estimates for $\alpha(1)$ is not robust as the probability interval lies within the negative and positive region. This last result is in line with what we find in the previous sections. Regarding the TVTP model, we label Regime 1 as the "negatively skewed" Regime and Regime 2 as the "positively skewed" Regime. We let unchanged the label for the FTP model.

Regarding the posterior moments of γ_0 and γ_1 in the TVTP model, the posterior estimates for γ_0 and γ_1 are both positive, although the uncertainty surrounding the latter is relatively large. Regarding the parameters that govern the time-varying transition probabilities, the estimates for b_1 is vigorously negative, suggesting that an increase in the leading indicator (NFCI) increases the probability of switching from the positively-skewed regime to the negatively-skewed regime. By contrast, although the estimates for b_0 is negative (-0.5891), its Bayesian posterior probability intervals are relatively large, suggesting that b_0 could be dropped from the model.

Figure 6 reports the probability (at the mode) of moving from the right skewed regime to the left skewed regime as a function of financial conditions. For example, when the NFCI is below -1.5, the probability of moving to the regime of downside risk is close to zero, and thus the probability of remaining in the positive skewed regime is close to one. However, when NFCI rises, the probability of switching to the regime of downside risk starts increasing. More specifically, if the NFCI exceeds 1.50, then the probability of moving to the downside risk Regime becomes higher than 50 percent.

To provide a full characterization of downside risks of GDP growth, Figure 7 plots the fitted conditional probability density functions of GDP growth across TVTP and FTP models for six sample periods at different points of the business cycle[13] : 1973.Q1, which represents the period before the 1973 oil crisis; 1974.Q4 which is the first quarter of the beginning of a NBER recession; 1978.Q4, which represents the end of an expansion phase before the second oil shock; 1981.Q2, which represents a change in the Federal Reserve's operating procedures; 2006.Q2, which marked the end of the Federal Reserve tightening before the financial crisis; and 2008.Q4 which represents the period after the failure of Lehman. The figure reveals important insights into how much the density function of GDP growth varies from the two specifications. During expansionary phases such as in 1973.Q1, 1978.Q4 and 2006.Q2, conditional densities are remarkably similar across both models; variances are relatively low and skewnesses tends to be positive. During downturns, such as 2008.Q4, conditional densities differ considerably. FTP densities do not reveal any form of skewness, while TVTP densities tend to be negatively skewed as they cover much wider negative range.

Overall, results like this reveals that the information content of financial conditions for downside risks of GDP is particularly strong. By providing valuable additional information

[13]To characterize the risk of GDP growth, we construct means of predictive density of GDP growth as follows:

$$p(y_t|Y_{t-1},\theta) = \sum_{k=1}^{2} \Pr(s_t = k)\frac{2}{\sigma}\phi\left(\frac{y_t - \phi x_t}{\sigma}\right)\Phi\left(\alpha(k)\frac{y_t - \phi x_t}{\sigma}\right), \tag{48}$$

where $\Pr(s_t = k)$ is the filtered probabilities of being in regime k at time t.

about whether a particular regime has occurred, financial conditions turn out to be able to capture and predict downside risks to GDP growth. Our results are thus consistent with macro-finance theory.

VIII. Conclusion

Our main goal in this paper was to provide a characterization of cyclical variations in the risk of GDP growth that go beyond the first few moments. To do so, we developed a Gibbs sampling procedure for autoregressive time series subject to Markov skewness shifts. An empirical application of this methodology to postwar U.S. GDP has shown that, in periods of expansion, the GDP distribution tends to be asymmetric with a higher weight on positive outcomes, due to a large positive skewness. By contrast, the distribution tends to be symmetric in periods of recession. These findings are robust to various alternative specifications.

A straightforward extension of our approach base led us to naturally test, in a Bayesian manner, the nonlinear relationship between GDP skewness and financial conditions. We extended the model base to include a measure of financial conditions that is able to shift the transition probabilities. While existing macro-finance theories document linkages between downside risks to GDP growth (i.e., negative skewness) and tighter financial conditions, the empirical evidence has been scant. We provided strong evidence for these theories. We showed financial conditions help to identify a regime characterized by downside risks, while the fixed transition probability model cannot capture this behavior.

Extending univariate AR models with Markov skewness shifts to a multivariate framework, like VAR models, would seem to be a natural next step. Another area of future work would be to relax the assumption of exogeneity of regime switching in order to better understand the sources of changes in the skewness of GDP growth. As such, the works by Kim, Piger, and Startz (2008) and Chang, Choi, and Park (2017) on endogenous Markov-switching AR models could then be used in this direction. All in all, we 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 (forthcoming): "Vulnerable Growth," *American Economic Review.*

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.

Baker, S. R., N. Bloom, and S. J. Davis (2016): “Measuring Economic Policy Uncertainty,” *The Quarterly Journal of Economics*, 131(4), 1593–1636.

Barro, R. J. (2006): “Rare Disasters and Asset Markets in the Twentieth Century,” *The Quarterly Journal of Economics*, 121(3), 823–866.

——— (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.

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

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

Bloom, N., F. Guvenen, and S. Salgado (2016): “Skewed Business Cycles,” 2016 Meeting Papers 1621, Society for Economic Dynamics.

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

Burns, A. F., and W. C. Mitchell (1946): *Measuring Business Cycles*. National Bureau of Economic Research, Inc.

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.

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.

Clark, T. E., A. Carriero, and M. Massimiliano (2017): "Measuring Uncertainty and Its Impact on the Economy," *The Review of Economics and Statistics*.

Diebold, F., J.-H. Lee, and G. . Weinbach (1994): "Regime Switching with Time-Varying Transition Probabilities," *Nonstationary Time Series Analysis and Cointegration*, pp. 283–302.

Diebold, F. X., and G. D. Rudebusch (1996): "Measuring Business Cycles: A Modern Perspective," *The Review of Economics and Statistics*, 78(1), 67–77.

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

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

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.

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.

——— (2016): "Chapter 3 - Macroeconomic Regimes and Regime Shifts," vol. 2 of *Handbook of Macroeconomics*, pp. 163 – 201. Elsevier.

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.

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

Keynes, J. M. (1936): *The General Theory of Employment, Interest and Money*. Macmillan, 14th edition, 1973.

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

Kim, C.-J., and C. R. Nelson (1999): *State-space Models with Regime Switching*, MIT Press Books. The MIT Press.

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.

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 (2016): "Do Uncertainty Shocks Always Matter for Business Cycles?," Working Papers 2016-19, CEPII research center.

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.

McConnell, M. M., and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed since the Early 1980's?," *American Economic Review*, 90(5), 1464–1476.

Mitchell, W. C. (1927): *Business Cycles: The Problem and Its Setting*. National Bureau of Economic Research, Inc.

Morley, J., and J. Piger (2012): "The Asymmetric Business Cycle," *The Review of Economics and Statistics*, 94(1), 208–221.

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

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

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

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

Stock, J. H., and M. W. Watson (2003): "Has the Business Cycle Changed and Why?," in *NBER Macroeconomics Annual 2002, Volume 17*, NBER Chapters, pp. 159–230. National Bureau of Economic Research, Inc.

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.

Yellen, J. L. (2016): "The Outlook, Uncertainty, and Monetary Policy, Remarks at the Economic Club of New York, New York, New York," https://www.federalreserve.gov/newsevents/speech/yellen20160329a.htm.

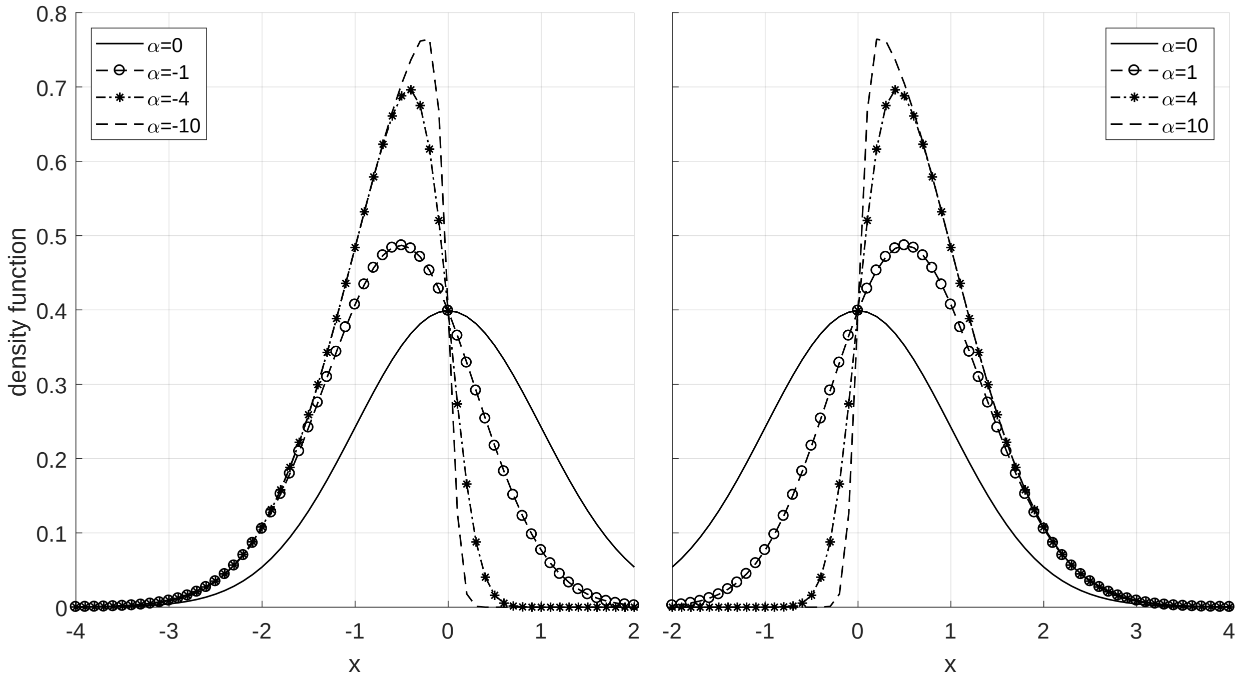

Figure 1. Skew-normal density functions when $\alpha = 0, -1, -4, -10$ in the left-hand panel, and $\alpha = 0, 1, 4, 10$ in the right-hand panel. The location parameter, ξ, is set to zero, and the scale parameter, σ^2 to one.

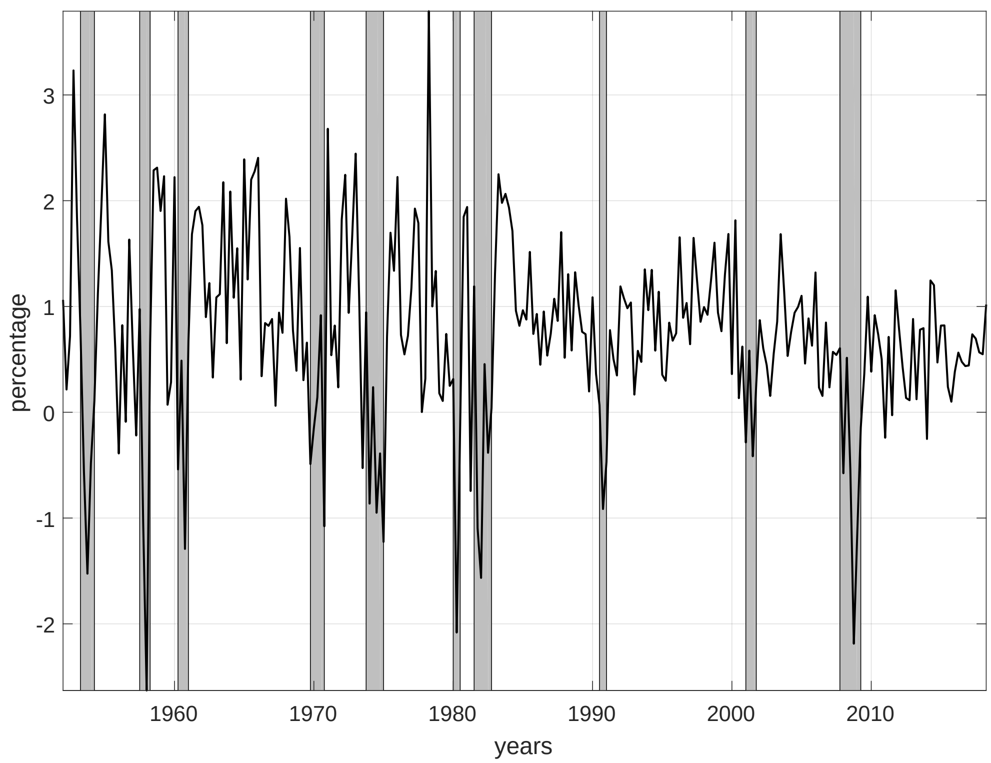

Figure 2. Sample period: 1952.Q1 — 2018.Q2. Quarterly growth rates of postwar U.S. real GDP. The grey areas denote the NBER recessions.

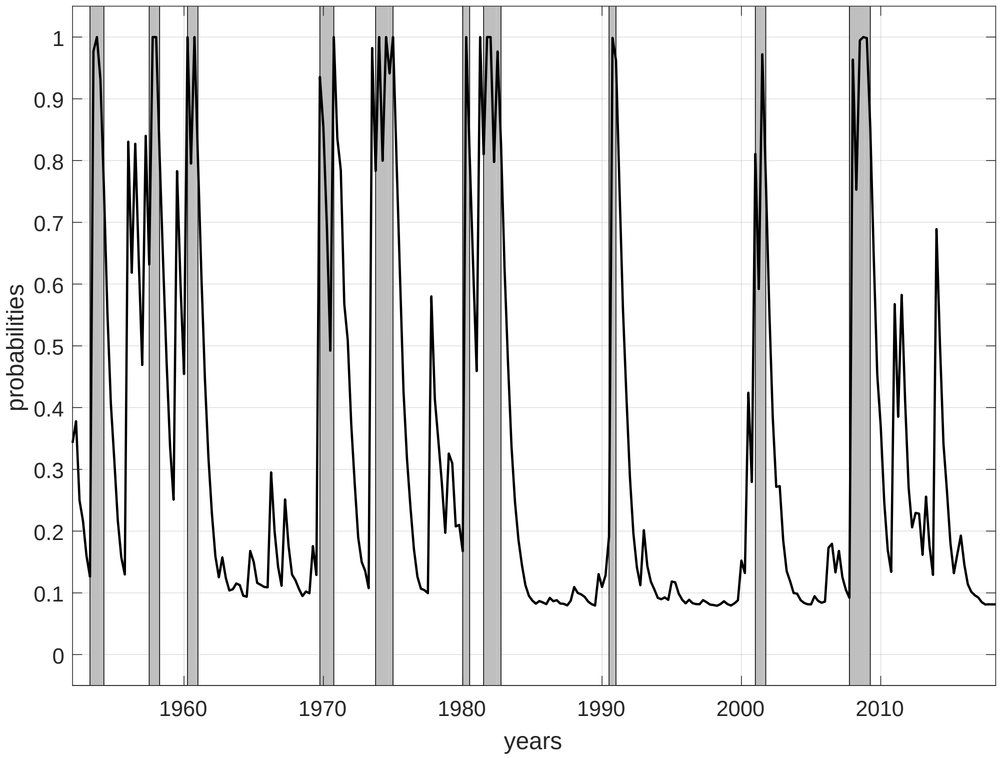

Figure 3. Sample period: 1952.Q1 — 2018.Q2. Filtered probabilities of Regime 1 (i.e., the symmetric regime). The grey areas denote the NBER recessions.

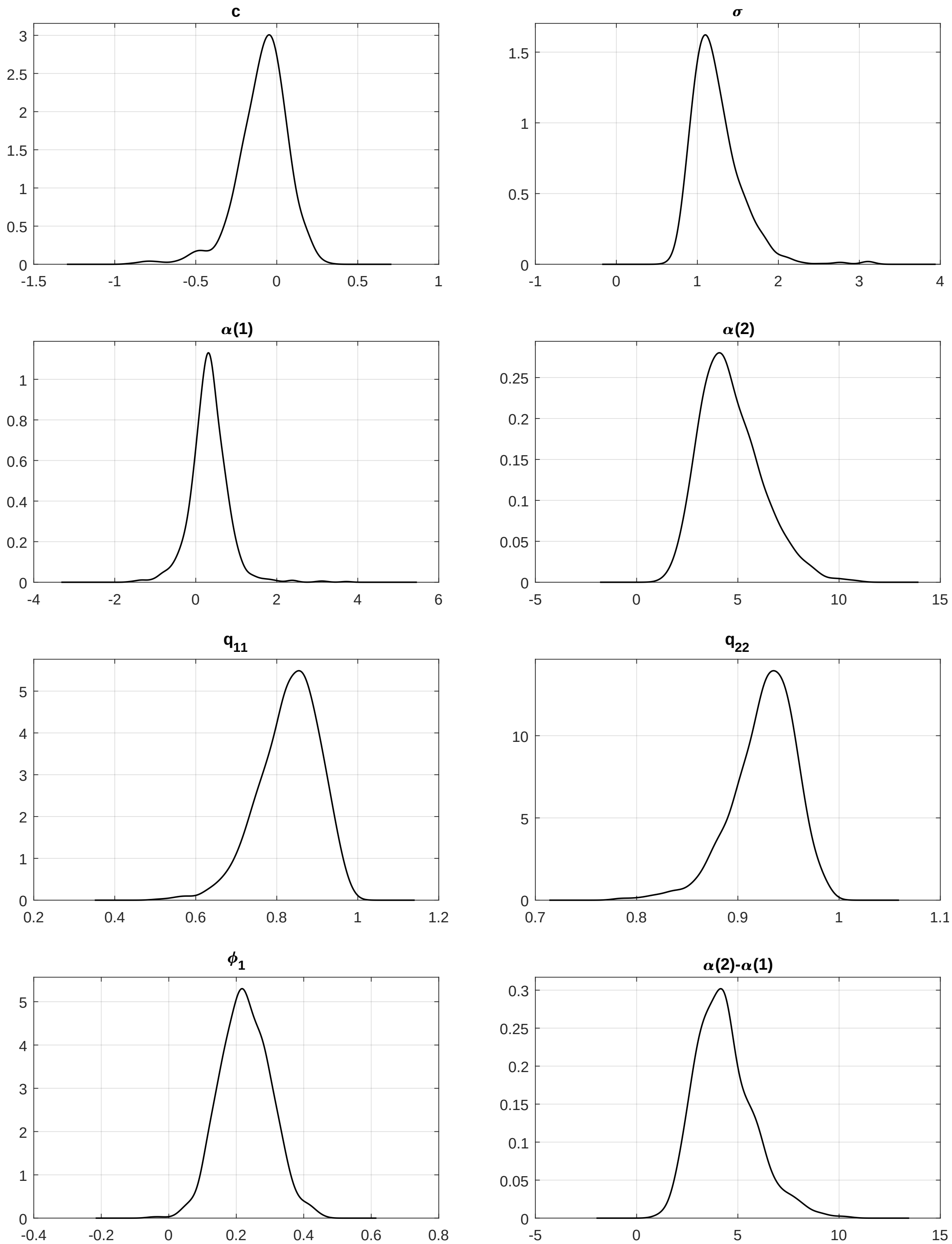

FIGURE 4. Marginal posterior densities using normal kernel density estimates produced from the AR(1) model with Markov skewness shifts.

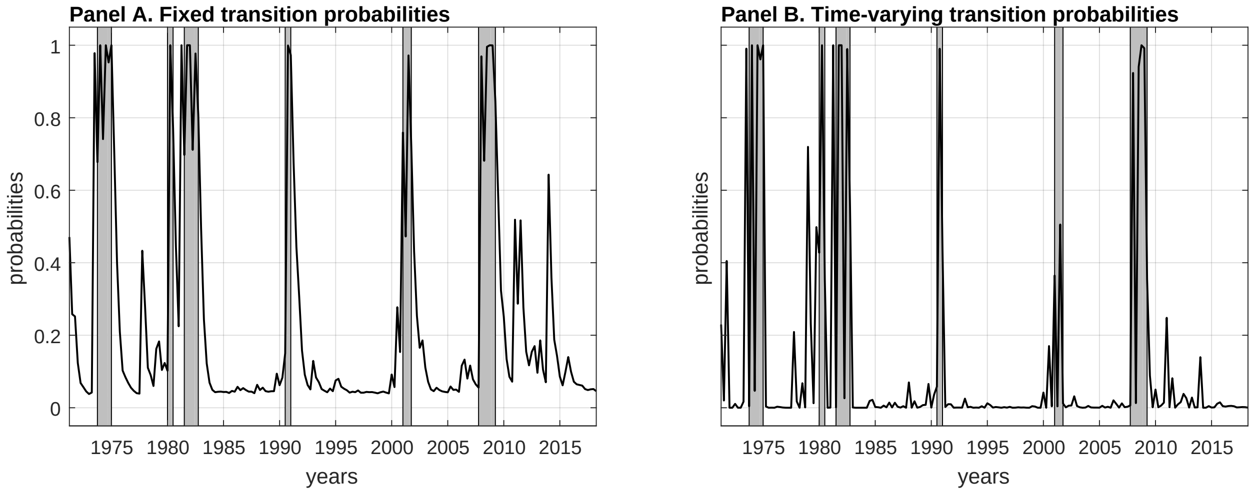

FIGURE 5. Sample period: 1971.Q2 — 2018.Q2. Probabilities (at the mode) of being in Regime 1 produced from the model with fixed transition probabilities (Panel A; i.e, left-hand side panel) and the model with time-varying transition probabilities (Panel B; i.e, the right-hand side panel). The grey areas denote the NBER recessions.

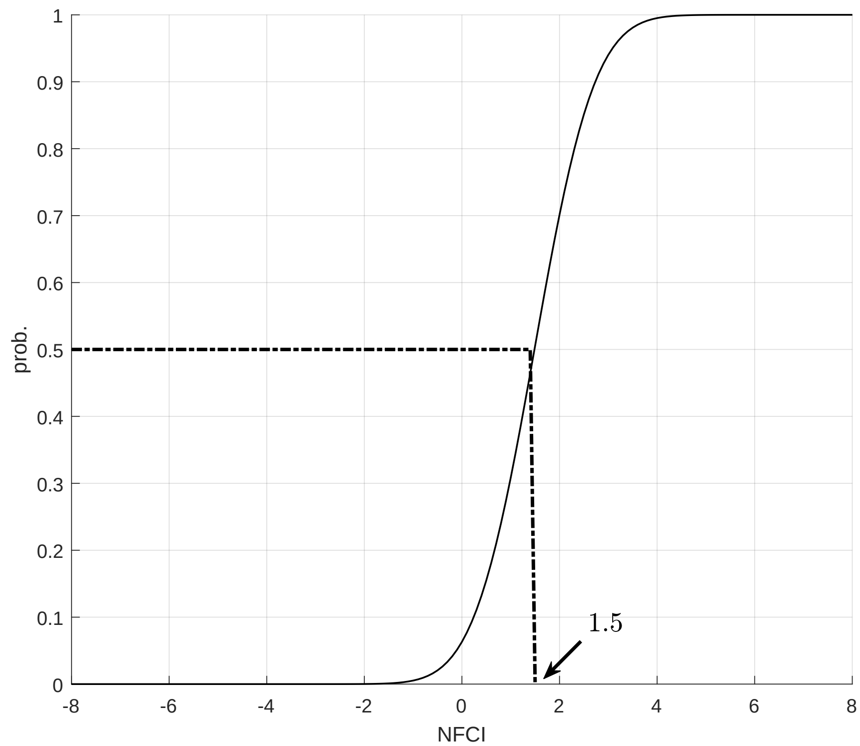

FIGURE 6. Probability (at the mode) of moving from the right skewed regime to the left skewed regime as a function of financial conditions.

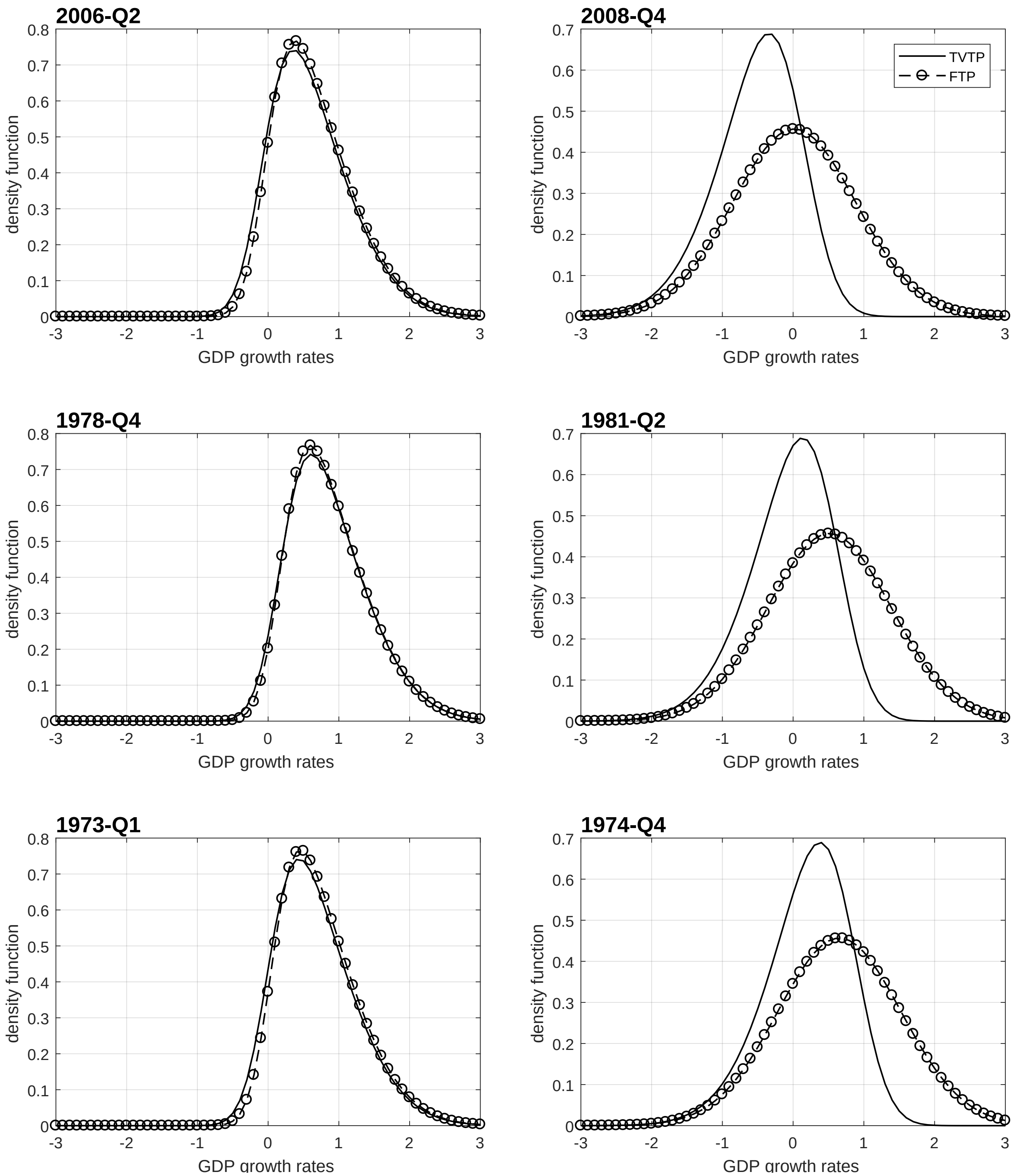

FIGURE 7. Density functions for real GDP growth rates, produced from the model with fixed transition probabilities (dotted black line) and the model with time-varying transition probabilities (solid black line), at different horizons: 2008-Q4, 2006-Q2, 1981-Q2, 1978-Q4, 1974-Q4, and 1973-Q1.

TABLE 1. AR(1) Markov-switching model for U.S. real GDP growth rates.

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
c	constant	N	0.00	5.00	0.0457	−0.0657	−0.0506	−0.3190	0.1347
σ	scale	Inv-G	1.00	1.00	0.9943	1.2071	1.1533	0.8535	1.7182
$\alpha(1)$	shape	N	0.00	3.00	0.1496	0.2421	0.2472	−0.5014	0.9604
$\alpha(2)$	shape	N	0.00	3.00	3.5034	4.6479	4.4758	2.6052	7.4478
q_{11}	prob.	B	0.80	0.10	0.8365	0.8032	0.8121	0.8440	0.9146
q_{22}	prob.	B	0.80	0.10	0.9175	0.9094	0.9147	0.8440	0.9560
ϕ_1	persistence	N	0.00	5.00	0.2140	0.2210	0.2210	0.1029	0.3489
$\alpha(2)-\alpha(1)$	shape	-	-	-	3.3538	4.4058	4.2217	2.4921	7.1597

Note: Sample period: 1952.Q1—2018.Q2. N stands for Normal, B for Beta, and Inv-G for Inverted-Gamma distributions. 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.

TABLE 2. Information criteria

			difference relative to AR(1)	
Model	WAIC	WAIC standard errors	WAIC	WAIC standard errors
Skewness shifts AR(1)	303.0985	12.6483	0.0000	0.0000
Skewness shifts AR(0)	301.0800	13.6789	2.0185	3.9433
Skewness shifts AR(2)	304.4611	12.8252	−1.3626	2.2059
Skewness shifts AR(3)	303.2791	13.2823	−0.1806	2.8681
Skewness shifts AR(4)	304.1774	13.5933	−1.0789	3.2674
Constant normal AR(1)	332.1437	15.6100	−29.0452	6.7203
Constant skew-normal AR(1)	333.6921	15.9603	−30.5935	6.8243
MS Hamilton AR(1)	328.4164	15.9091	−25.3179	6.3785

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 $se(\overline{elpd}^{A} - \overline{elpd}^{B}) = \sqrt{T * V_{t=1}^{T}(\overline{elpd,t}^{A} - \overline{elpd,t}^{B})}$.

TABLE 3. AR(1) Markov-switching model for U.S. real GDP growth rates with mean and skewness drifts.

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
$c(1)$	constant	N	0.00	5.00	0.3581	0.4312	0.5965	-0.7421	1.1531
$c(2)$	constant	N	0.00	5.00	-0.0105	0.0786	0.0521	-0.1605	0.3850
σ	scale	Inv-G	1.00	1.00	0.9738	1.1112	1.0866	0.9107	1.3698
$\alpha(1)$	shape	N	0.00	3.00	0.0007	-1.3777	-0.7306	-5.1699	1.0876
$\alpha(2)$	shape	N	0.00	3.00	3.6480	4.2200	4.1166	1.7861	7.1306
q_{11}	prob.	B	0.80	0.10	0.8884	0.7863	0.7940	0.6499	0.8990
q_{22}	prob.	B	0.80	0.10	0.9194	0.8739	0.8891	0.7546	0.9456
ϕ_1	persistence	N	0.00	5.00	0.2603	0.1799	0.1789	0.0669	0.2949
$\alpha(2) - \alpha(1)$	shape	-	-	-	3.6473	5.5977	5.1023	2.6046	10.3006

Note: Sample period: 1952.Q1—2018.Q2. N stands for Normal, B for Beta, and Inv-G for Inverted-Gamma distributions. 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.

TABLE 4. AR(1) Markov-switching model for U.S. real GDP growth rates throughout the pre-Great Moderation period.

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
c	constant	N	0.00	5.00	0.4083	0.5426	0.4546	−0.1037	1.5628
σ	scale	Inv-G	1.00	1.00	1.0907	1.1839	1.1482	0.9700	1.5004
$\alpha(1)$	shape	N	0.00	3.00	−0.3263	−0.8684	−0.5265	−3.7022	0.5295
$\alpha(2)$	shape	N	0.00	3.00	1.5721	1.5076	1.5606	−1.0521	4.1486
q_{11}	prob.	B	0.80	0.10	0.8194	0.7864	0.7945	0.6277	0.9200
q_{22}	prob.	B	0.80	0.10	0.8579	0.8027	0.8191	0.6319	0.9214
ϕ_1	persistence	N	0.00	5.00	0.1745	0.2158	0.2234	0.0003	0.4056
$\alpha(2)-\alpha(1)$	shape	-	-	-	1.8984	2.3759	2.2033	0.2790	4.9478

Note: Sample period: 1952.Q1—1984.Q4. N stands for Normal, B for Beta, and Inv-G for Inverted-Gamma distributions. 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.

TABLE 5. AR(1) Markov-switching model for U.S. real GDP growth rates throughout the post-Great Moderation period.

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
c	constant	N	0.00	5.00	0.1961	−0.0074	0.0402	−0.4758	0.2739
σ	scale	Inv-G	1.00	1.00	0.6210	1.0533	0.8707	0.5566	2.2038
$\alpha(1)$	shape	N	0.00	3.00	−0.0029	−0.0718	−0.0140	−2.3154	1.9290
$\alpha(2)$	shape	N	0.00	3.00	2.5912	4.3480	4.1030	1.8946	7.8486
q_{11}	prob.	B	0.80	0.10	0.8300	0.7572	0.7671	0.6006	0.8880
q_{22}	prob.	B	0.80	0.10	0.9159	0.9106	0.9226	0.8085	0.9738
ϕ_1	persistence	N	0.00	5.00	0.2127	0.2129	0.2100	−0.0080	0.4483
$\alpha(2)-\alpha(1)$	shape	-	-	-	2.5941	4.4198	4.0871	1.6313	8.3641

Note: Sample period: 1985.Q1—2018.Q4. N stands for Normal, B for Beta, and Inv-G for Inverted-Gamma distributions. 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.

Table 6. AR(1) Markov-switching model for U.S. real GDP growth rates in the presence of time-varying transition probabilities (as a function of financial conditions)

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
c	constant	N	0.00	5.00	0.0548	0.0230	0.0312	−0.1553	0.1829
σ	scale	Inv-G	1.00	1.00	0.8762	0.9338	0.9177	0.7219	1.2135
$\alpha(1)$	shape	N	0.00	3.00	−1.8931	−1.9716	−1.5878	−5.4079	−0.0200
$\alpha(2)$	shape	N	0.00	3.00	2.7849	3.0965	2.9274	1.7168	5.1323
ϕ_1	persistence	N	0.00	5.00	0.2066	0.2125	0.2135	0.0958	0.3264
γ_0	prob.	N	0.00	2.00	1.3785	1.3572	1.3048	−0.0273	3.0662
γ_1	prob.	N	0.00	2.00	0.1617	0.2914	0.2781	−1.3329	2.0057
b_0	prob.	N	0.00	2.00	−0.5891	−0.7005	−0.6567	−1.9243	0.3487
b_1	prob.	N	0.00	2.00	−1.0394	−1.3888	−1.2305	−2.7597	−0.6818
$\alpha(2) - \alpha(1)$	shape	-	-	-	4.6780	5.0681	4.6764	2.6402	8.7362

Note: Sample period: 1971.Q2—2018.Q2. N stands for Normal, and Inv-G for Inverted-Gamma distributions. 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.

Table 7. AR(1) Markov-switching model for U.S. real GDP growth rates in the presence of fixed transition probabilities

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
c	constant	N	0.00	5.00	0.0849	−0.0110	0.0041	−0.2448	0.1720
σ	scale	Inv-G	1.00	1.00	0.8693	1.1119	1.0533	0.7333	1.7002
$\alpha(1)$	shape	N	0.00	3.00	−0.1093	−0.0805	−0.0564	−0.8850	0.6441
$\alpha(2)$	shape	N	0.00	3.00	3.2600	4.7520	4.5754	2.5061	7.7491
q_{11}	prob.	B	0.80	0.10	0.8102	0.7789	0.7876	0.6364	0.8960
q_{22}	prob.	B	0.80	0.10	0.9312	0.9216	0.9276	0.8572	0.9654
ϕ_1	persistence	N	0.00	5.00	0.1877	0.1869	0.1843	0.0570	0.3160
$\alpha(2) - \alpha(1)$	shape	-	-	-	3.3693	4.8325	4.6456	2.5961	7.9368

Note: Sample period: 1971.Q2—2018.Q2. N stands for Normal, B for Beta, and Inv-G for Inverted-Gamma distributions. 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.

Appendix A. The likelihood, $p(Y_T|\theta)$

The evaluation of the overall likelihood function is obtained using the standard Hamilton (1989) filter. The likelihood of Y_T is

$$p(Y_T|\theta) = \prod_{t=1}^{T} p(y_t|Y_{t-1}, \theta), \tag{49}$$

where the conditional likelihood function $p\left(y_t|Y_{t-1}, \theta\right)$, given θ, at date t is obtained by integrating the density $p(y_t, s_t|Y_{t-1}, \theta)$ over s_t as follows

$$p(y_t|Y_{t-1}, \theta) = \sum_{s_t \in H} p(y_t, s_t|Y_{t-1}, \theta), \tag{50}$$

$$= \sum_{s_t \in H} p(y_t|s_t, Y_{t-1}, \theta)\text{Pr}[s_t|Y_{t-1}, \theta], \tag{51}$$

Using the Hamilton (1989) filter, we can recursively compute $\text{Pr}[s_t|Y_t, \theta]$ forward. Specifically,

$$\text{Pr}[s_t|Y_{t-1}, \theta] = \sum_{s_{t-1} \in H} q_{s_t, s_{t-1}} \text{Pr}(s_{t-1}|Y_{t-1}, \theta), \qquad \text{for} \quad t > 0, \tag{52}$$

where $q_{s_t, s_{t-1}} = \text{Pr}[s_t|s_{t-1}]$ is the transition probability described in (8).

We then update the joint probability term in the following way:

$$\text{Pr}[s_t|Y_t, \theta] = \frac{p(y_t, s_t|Y_{t-1}, \theta)}{p(y_t|Y_{t-1}, \theta)} \tag{53}$$

$$= \frac{p(y_t|s_t, Y_{t-1}, \theta).\text{Pr}(s_t|Y_{t-1}, \theta)}{p(y_t|Y_{t-1}, \theta)}, \qquad \text{for} \quad t > 0, \tag{54}$$

Once the parameters of the model are estimated, we follow Kim (1994) and Kim and Nelson (1999) and make inference on s_T, the smoothed probabilities, in the following way:

$$\text{Pr}[s_t|Y_T, \theta] = \sum_{s_{t+1} \in H} \text{Pr}[s_t, s_{t+1}|Y_T, \theta], \tag{55}$$

where

$$\text{Pr}[s_t, s_{t+1}|Y_T, \theta] = \frac{\text{Pr}[s_{t+1}|Y_T, \theta].\text{Pr}[s_t|Y_T, \theta].\text{Pr}[s_{t+1}|s_t]}{\text{Pr}[s_{t+1}|Y_T, \theta]}. \tag{56}$$

The advantage of such a method is that it allows us to infer the unobservable variable s_t using all the information in the sample.

Appendix B. Computer Software

Once compiled, our companion C++ computer code for this paper, available at the author's website, is easy to use. One must provide an input file that indicates prior specifications, the structure of AR process, MCMC options, and time series data. An example of this interface is provided below.

```
//== Number Lagged Variables ==//
1

//== Size Sample ==//
267

//== Number Regimes ==//
2

//== Prior for Phi ==//
0.0000   5.0000

//== Prior for Scale ==//
1.1782   2.5891

//== Prior for Shape ==//
0.0000   3.0000

//== Prior for Transition Matrix ==//
12.00   3.000
3.000   12.00

//== Number Optimization Runs ==//
100

//== Number Draws ==//
11000

//== Number Burn-in ==//
1000

//== Thinning Factor ==//
10

```

```
//== Y-data ==//
2.1999843197342273e-01
1.0628799925268773e+00
...
```

This input file concerns an AR model of order 1 where the shape parameter follows a two-states Markov process. Regarding the MCMC procedure, the input file asks for $11,000$ draws, whose $1,000$ as burn-in, and 10 as thinning factor. In this file, the header bracketed by

```
//== ... ==//
```

communicates with the software what kind of data is expected. The number below the header "Number Lagged Variables" indicates how many lags are defined for the model. The number below the header "Size Sample" indicates the size of the sample. The number below the header "Number Regimes" indicates the number of regimes for the Markov process. The number below the header "Prior for Phi", "Prior for Scale", "Prior for Shape", and "Prior for Transition Matrix" indicate the hyperparameters for the parameters ϕ, σ, α, and Q, respectively. The number below the header "Number Optimization Runs" indicates the number of times the optimization process is repeated. At each time, a new set of points, generated from the prior, is used. The number below the header "Number Draws" indicates the number of draws in the MCMC algorithm. The number below the header "Number Burn-in" indicates the number of burn-in in the MCMC algorithm. The number below the header "Thinning Factor" indicates the thinning factor in the MCMC algorithm. Finally, the values below the header "Y-data" contain time series data.

Appendix C. Additional Figures

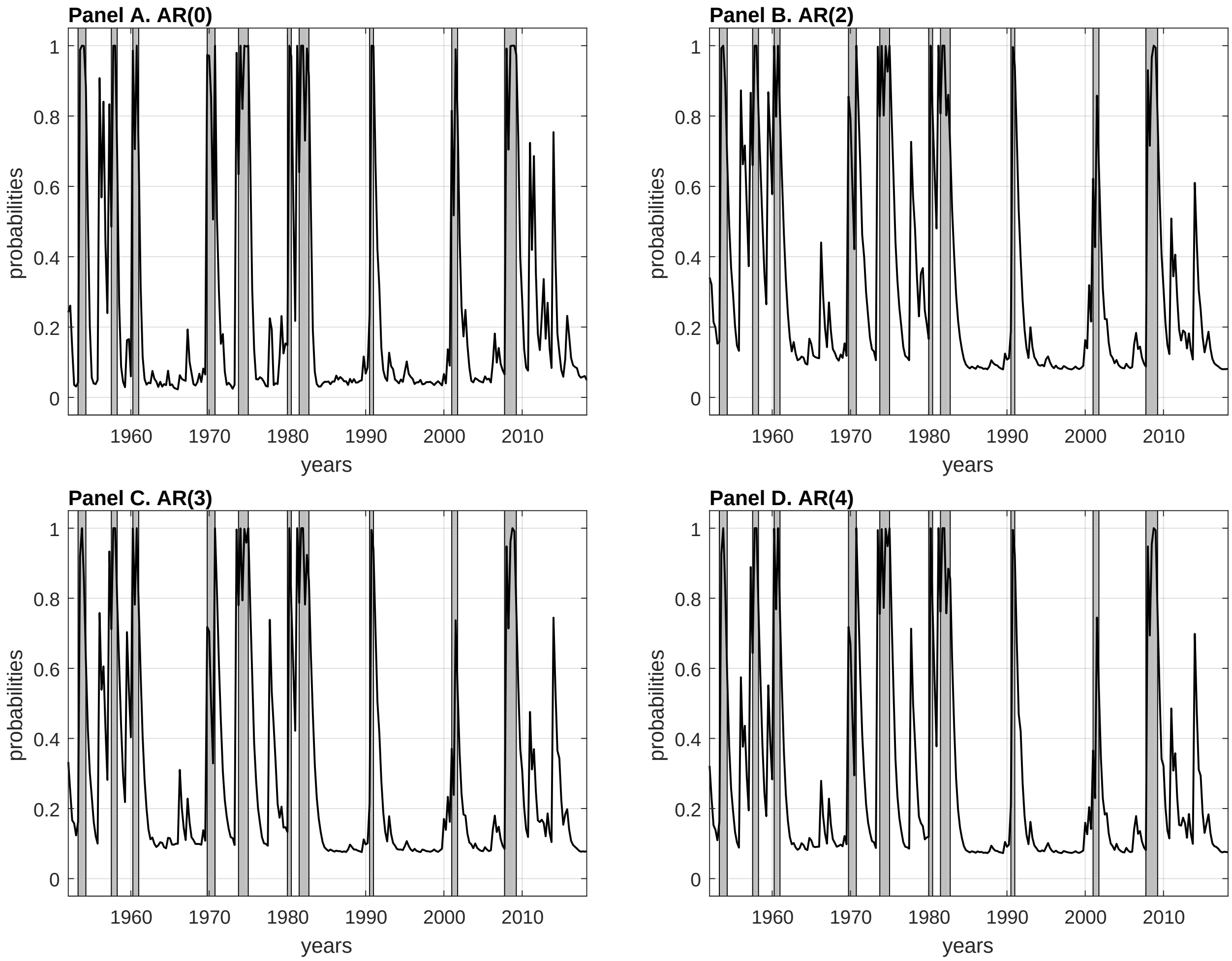

Figure 8. Sample period: 1952.Q1 — 2018.Q2. Filtered probabilities of the non-skewed regime across alternative autoregressive specifications. The grey areas denote the NBER recessions.

Appendix D. Additional Tables

Table 8. AR(0) Markov-switching model for U.S. real GDP growth rates.

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
c	constant	N	0.00	5.00	0.2411	0.1962	0.1936	0.0127	0.3631
σ	scale	Inv-G	1.00	1.00	1.0143	1.1554	1.0908	0.7956	1.7299
$\alpha(1)$	shape	N	0.00	3.00	−0.3669	−0.4608	−0.3834	−1.3808	0.2124
$\alpha(2)$	shape	N	0.00	3.00	3.7625	4.6085	4.3556	2.4263	7.5978
q_{11}	prob.	B	0.80	0.10	0.8021	0.7777	0.7838	0.6544	0.8823
q_{22}	prob.	B	0.80	0.10	0.9240	0.9151	0.9175	0.8667	0.9530
$\alpha(2)-\alpha(1)$	shape	-	-	-	4.1294	5.0693	4.8292	2.9333	8.0641

Note: Sample period: 1952.Q1—2018.Q2. N stands for Normal, B for Beta, and Inv-G for Inverted-Gamma distributions. 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.

Table 9. AR(2) Markov-switching model for U.S. real GDP growth rates.

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
c	constant	N	0.00	5.00	−0.0389	−0.2021	−0.1720	−0.5295	0.0460
σ	scale	Inv-G	1.00	1.00	0.9960	1.2892	1.2062	0.8959	1.8883
$\alpha(1)$	shape	N	0.00	3.00	0.2014	0.4579	0.3779	−0.2314	1.3706
$\alpha(2)$	shape	N	0.00	3.00	3.5441	4.9742	4.7294	2.7788	7.8311
q_{11}	prob.	B	0.80	0.10	0.8322	0.7977	0.8077	0.6532	0.9067
q_{22}	prob.	B	0.80	0.10	0.9162	0.9037	0.9119	0.8303	0.9560
ϕ_1	persistence	N	0.00	5.00	0.1952	0.2073	0.2094	0.0862	0.3267
ϕ_2	persistence	N	0.00	5.00	0.1129	0.1313	0.1306	0.0231	0.2441
$\alpha(2)-\alpha(1)$	shape	-	-	-	3.3426	4.5163	4.3030	2.5464	7.1865

Note: Sample period: 1952.Q1—2018.Q2. N stands for Normal, B for Beta, and Inv-G for Inverted-Gamma distributions. 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.

Table 10. AR(3) Markov-switching model for U.S. real GDP growth rates.

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
c	constant	N	0.00	5.00	0.0101	−0.1378	−0.1212	−0.4164	0.1033
σ	scale	Inv-G	1.00	1.00	0.9969	1.2506	1.1897	0.8673	1.8152
$\alpha(1)$	shape	N	0.00	3.00	0.2328	0.3645	0.3402	−0.2851	0.9772
$\alpha(2)$	shape	N	0.00	3.00	5.2637	4.9787	4.8044	2.7460	7.8610
q_{11}	prob.	B	0.80	0.10	0.8221	0.7968	0.8062	0.6642	0.9012
q_{22}	prob.	B	0.80	0.10	0.9004	0.9085	0.9143	0.8419	0.9577
ϕ_1	persistence	N	0.00	5.00	0.2088	0.2304	0.2281	0.1160	0.3447
ϕ_2	persistence	N	0.00	5.00	0.1347	0.1588	0.1576	0.0474	0.2750
ϕ_3	persistence	N	0.00	5.00	−0.1067	−0.1195	−0.1159	−0.2201	−0.0260
$\alpha(2)-\alpha(1)$	shape	-	-	-	5.0309	4.6141	4.4386	2.5865	7.3546

Note: Sample period: 1952.Q1—2018.Q2. N stands for Normal, B for Beta, and Inv-G for Inverted-Gamma distributions. 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.

Table 11. AR(4) Markov-switching model for U.S. real GDP growth rates.

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
c	constant	N	0.00	5.00	0.0379	−0.0987	−0.0873	−0.3795	0.1557
σ	scale	Inv-G	1.00	1.00	0.9961	1.2443	1.1856	0.8819	1.7981
$\alpha(1)$	shape	N	0.00	3.00	0.1282	0.3172	0.3087	−0.4077	1.0374
$\alpha(2)$	shape	N	0.00	3.00	3.5330	4.8866	4.6920	2.7498	7.6820
q_{11}	prob.	B	0.80	0.10	0.8108	0.7865	0.7931	0.6549	0.8987
q_{22}	prob.	B	0.80	0.10	0.9188	0.9081	0.9135	0.8424	0.9568
ϕ_1	persistence	N	0.00	5.00	0.2049	0.2211	0.2209	0.0976	0.3408
ϕ_2	persistence	N	0.00	5.00	0.1473	0.1619	0.1598	0.0601	0.2748
ϕ_3	persistence	N	0.00	5.00	−0.1019	−0.1123	−0.1100	−0.2105	−0.0194
ϕ_4	persistence	N	0.00	5.00	−0.0511	−0.0485	−0.0474	−0.1510	0.0589
$\alpha(2)-\alpha(1)$	shape	-	-	-	3.4048	4.5694	4.3319	2.5474	7.2284

Note: Sample period: 1952.Q1—2018.Q2. N stands for Normal, B for Beta, and Inv-G for Inverted-Gamma distributions. 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.