THE SWITCHING SKEWNESS OVER THE BUSINESS CYCLE

STÉPHANE LHUISSIER*

Abstract. Motivated by the analysis of the evolution of the distribution of macroeconomic time series data over time, this paper develops and applies a Gibbs-sampler for autoregressive time series subject to regime switches in the tails of the distribution. More specifically, we consider the sknew-normal distribution proposed by Azzalini (1985, 1986), in which the shape parameter is allowed to change over time according to a Markov-switching process. As an empirical illustration, we analyse the distribution of the growth rates of postwar U.S. real GDP, and find periodic shifts between a left and right-skewed distribution regime, with the former corresponding closely to NBER recession dates. Hence, more theorizing is needed to better understand the interaction between variation in tails and the business cycle.

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. From the use of simple descriptive statistics to the application of the most sophisticated time-varying coefficient models, 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. Indeed, recession phase is commonly associated with negative and volatile growth rates, while expansion phase is characterized by positive and stable growth rates.[1] In this paper we focus on the autoregressive (AR) model that is subject to third-moment shifts. Specifically, the skewness of the time series is governed by a Markov-switching process with unknown transition probabilities. Hamilton

Date: July 17, 2018.

Key words and phrases. Markov-switching, Skewness, Business cycle, Gibbs-sampler.

JEL Classification: C11; C2; E32.

*Banque de France, 31, Rue Croix des Petits Champs, DGEI-DEMFI-POMONE 41-1422, 75049 Paris Cedex 01, FRANCE (Email: stephane_lhuissier@club.fr; URL: 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.

[1]These findings result from a substantial body of the macroeconomic literature. See, for examples, Mitchell (1927), Keynes (1936), Burns and Mitchell (1946), Neftçi (1984), Hamilton (1989), and Diebold and Rudebusch (1996) on first-moment dynamics, and Bloom (2014) for a survey on second-moment dynamics.

(1989) uses a similar process to capture changes in the mean of the growth rates of postwar U.S. real Gross Domestic Product (GDP) 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 survey of this literature see Hamilton (2016).

The interest in the time-varying skewness of major business-cycle components stems from three different angles. First, if the behavior of skewness changes significantly between periods of expansion and recession, then macroeconomists should revise their theories and empirical macroeconomic models to take into account such a behavior and to explain it. Except for a few exceptions, we believe that there is no such model available within the profession. Our regime-switching framework can serve as the basis for describing the behavior of skewness over the business cycle in larger or more complicated models. Second, if the tails of the distribution change over the business cycle, then a model that generates such an asymmetry would deliver better predictive power for economic activity, a major asset for policymakers. Indeed, a complete empirical characterization of the distribution of future real GDP growth should paint a better picture of the state of the economy, and so lead to better suitable economic policies. Third, the presence of such asymmetry is essential to empirical work on Markov-switching rational expectations models. Those models include explicit expectations of future regime changes in the economy when solving the consumers and firms' optimization problems. If the behavior of skewness in major variables changes over the business cycle, then agents should take into account this possibility when forming expectations. As a consequence, regime shifts in the skewness of GDP distribution can have important effects on rational agents' expectation formation and on equilibrium dynamics.

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. Such a distribution has already been intensively studied in statistics, biologists, engineers, and medical researchers, but remains largely unexplored in economics.

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 skewness, in which one can generate in a flexible and straightforward manner alternatively draws from full conditional posterior distributions. The general theory of Gibbs sampling for finite mixture models of univariate skew-normal distributions has been developed by statisticians over the last decade (e.g., see for example Lin, Lee, Yen, and Chung, 2007; Frühwirth-Schnatter and Pyne, 2010). Our contribution to this existing theory pertains to developing a Gibbs sampler specifically for Markov mixture models, where regimes can exhibit persistence, and readily applicable to a class of widely used switching AR macroeconomic models.

Second, we apply our methodology to the growth rates of postwar U.S. real GDP. Our results show that the distribution of GDP growth exhibits large variation over time, where changes in the right and left tail probabilities are unambiguously strong. The times of changes are remarkably similar to NBER dating of business cycles. In particular, we find that recessions are associated with negative skewness — and, so, a left tail — while normal times are characterized by wider positive skewness — and, so, a right tail. It follows that, during recession periods, the probability of getting negative quartertly GDP growth rates is higher than observing positive GDP growth rates. By contrast, one observes the opposite phoenomenon during expansion periods, where the mass of the distribution is almost entirely concentrated on the positive region. The results are robusts when taking into account the possibility of changes in the mean and in the variance of GDP growth.

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 differences in economic dynamics across phases. The econometric work by Hamilton (1989) represents a major step forward, by modeling and testing this idea. The author allows the mean of the growth rates of U.S. Gross National Product (GNP) to vary over time according to a Markov-switching process. A limit employed in this paper is that the author assume a symmetric density function, and thus rule out, by construction, the possibility of skewed processes. By contrast, the flexibility of the skew-normal family employed in this paper seems particularly adecquate.

Much of the existing empirical literature has also analysed the second moment of GDP growth 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 relative to our paper lies within the fact that we depart from a symmetric distribution to emphasize the importance of asymmetric distribution.

More recently, there have been other studies that study 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 latter does not employ a Markov-switching framework.

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 Dynamic Stochastic General Equilibrium (DSGE) models. Our work differ from previous contributions in the sense that they all allow the mean and shock variances to vary over time, and thus rule out, by construction, the possibility of skewed distributions.

A growing number of theoretical and empirical studies have emphasized the importance of non-Gaussian shocks in macroeconomic models. See, for example, 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.

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 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 $\Phi(.)$ denotes 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

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

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

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

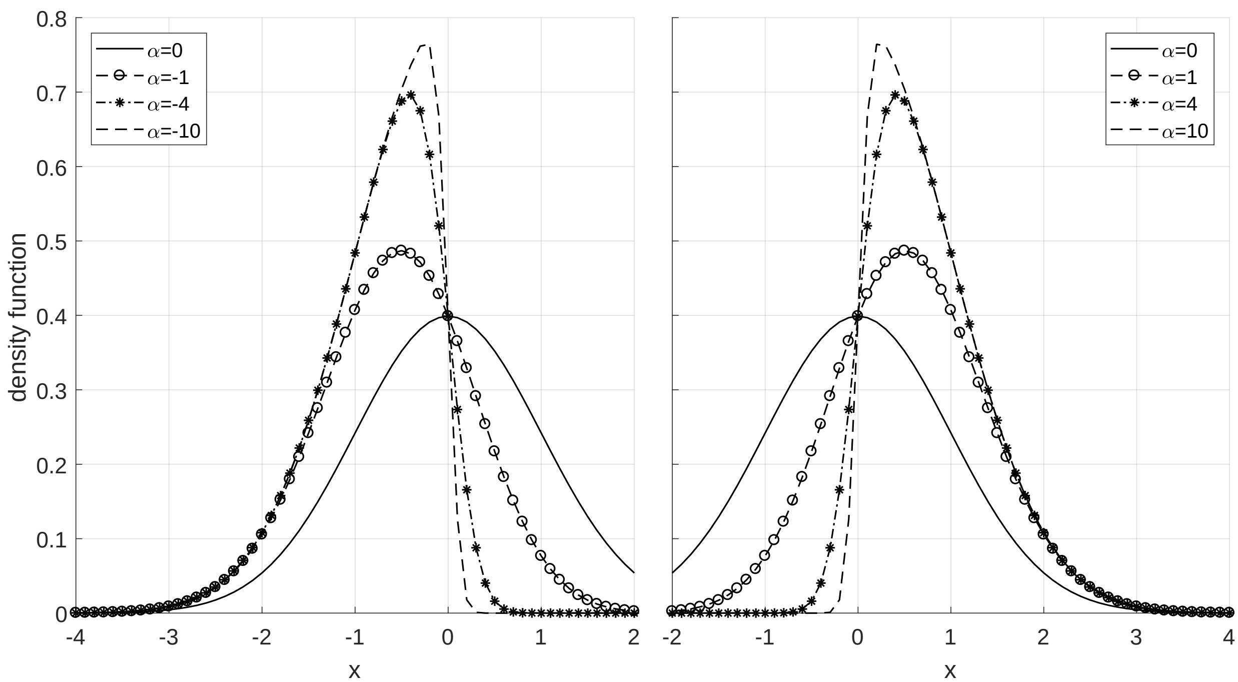

FIGURE 1. Skew-normal density functions when $\alpha = 0, -1, -4, -10$ in the left-hand panel, and $\alpha = 0, 1, 4, 10$ in the right-hand panel. The location parameter, ξ, is set to zero, and the scale parameter, σ^2 to one.

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 and variance Σ, 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. Markov-switching skewed autoregressive models

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

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

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 $\theta = (\phi, \sigma^2, \alpha, Q)$, where $\phi = [\phi_1, \ldots, \phi_\tau, c]$. 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 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 likelihood in (11) 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). However, 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. In this section, we introduce a Bayesian method to skew-normal models with Markov-switching. More specifically, we choose to extend Lin, Lee, Yen, and Chung (2007) as they make allowance for independently switches between the mean, the scale and the skewness, while Frühwirth-Schnatter and Pyne (2010) allow only switches in a synchronized manner.

Our approach can be also seen as an extension of Albert and Chib (1993) to Markov skew drifts. Indeed, the authors develop a Gibbs sampling for AR time series subject to only Markov mean and variance shifts.

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

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

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. Let $\phi = [...]$. The prior on the parameters θ and $\delta(k)$ is

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

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

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

$$\delta(k) = \text{uniform}(\delta(k)| -1, 1), \tag{20}$$

where $\bar{b}, \bar{B}, \bar{\alpha}, \bar{\beta}$ and $\bar{\alpha}_{1k}, \ldots, \bar{\alpha}_{hk}$ are the hyperparameters; uniform$(x|a, b)$ denotes the continuous uniform distribution in the interval $[a, b]$; 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 specify priors for the transformed parameters $\delta(k)$ introduced in Section II rather than directly for $\alpha(k)$ as we work directly with the stochastic representation.

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, and $Z_T = [z_1, \ldots, z_t]$. 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) $p\left(\delta^{(i)}|Y_T, S_T^{(i)}, Z_T^{(i)}, \theta_{\neq\delta}^{(i)}\right)$.

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 $p\left(\delta^{(i)}|Y_T, S_T^{(i)}, Z_T^{(i)}, \theta_{\neq\delta}^{(i)}\right)$ is not in closed form, and thus requires a Random-walk Metropolis-hasting (RWMH) algorithm to simulate draws from the posterior density.

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 non standard distribution of δ, where the implementation will be described below. The rest of the computation is straightforward as it is to sample from conditional densities in closed form. 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 $\delta(k)$, given Z_T, S_T and the remaining parameters.

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

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

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

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

where

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

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:

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

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

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

where

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

where $\mu_t^{(i)} = c^{(i)} + \phi_1^{(i)}(y_{t-1} - c^{(i)}) + \cdots + \phi_\tau{}^{(i)}(y_{t-\tau} - c^{(i)})$.

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

$$y_t^* = \phi x_t^* + \nu_t, \tag{28}$$

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

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

where

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

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

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

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

where ssr is the sum of squared residual defined as

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

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,** $p\left(\delta^{(i)}|Y_T, S_T^{(i)}, Z_T^{(i)}, \theta_{\neq\delta}^{(i)}\right)$**.** The full conditional distribution of δ is not of any standard form as shown below

$$p\left(\delta^{(i)}|Y_T, S_T^{(i)}, Z_T^{(i)}, \theta_{\neq\delta}^{(i)}\right) = \prod_{t=1}^{T}(1-\delta(s_t)^{(i)^2})^{\frac{1}{2}} \tag{34}$$

$$\times \exp\left\{-\frac{z_t^2 - 2\delta(s_t)^{(i)}z_t^{(i)}\left(y_t - \mu_t^{(i)}\right) + \left(y_t - \mu_t^{(i)}\right)^2}{2\sigma^2(1-\delta^2(s_t))}\right\}. \tag{35}$$

To simulate draws of δ, one can start to use a RWMH algorithm. To do so, we first transform $\delta(k)$, for $k = \{1, \ldots, H\}$, to $\delta^*(k) = \log\left\{\frac{1+\delta(k)}{1-\delta(k)}\right\}$, and then use the RWMH algorithm to the following density

$$g(\delta^*|Y_T, S_T, Z_T, \theta_{\neq\delta}) = p(\delta(\delta^*)|Y_T, S_T, Z_T, \theta_{\neq\delta})\prod_{k=1}^{H} J_{\delta^*(k)}, \tag{36}$$

where $J_{\delta^*(k)} = \frac{2\exp(\delta^*(k))}{(1+\exp(\delta^*(k)))^2}$ is the Jacobian of transformation from $\delta(k)$ to $\delta^*(k)$.

Here are the following steps of the algorithm

(1) Draw an $H \times 1$ vector, δ^p, from the following proposal distribution

$$\text{normal}(\delta^p|\delta^*, c^2\Sigma_{g^*}), \tag{37}$$

where Σ_{g^*} is the inverse of Hessian at the posterior mode of density (36) and c is the scale parameter.

(2) Calculate the acceptance probability as follows

$$\nu = \min\left\{1, \frac{g(\delta^p|Y_T, S_T, Z_T, \theta)}{g(\delta^{*(i-1)}|Y_T, S_T, Z_T, \theta)}\right\}, \tag{38}$$

where the conditional likelihood is evaluated using equation (36) according to the historical path of $S_T^{(i)}$.

(3) Draw a random number from an uniform distribution defined over the interval $[0, 1]$. If the generated number is less than or equal to the calculated value ν, we set $\delta^{*(i)} = \delta^p$. Otherwise, $\delta^{*(i)} = \delta^{*(i-1)}$. The rejection probability is chosen to be between 0.70 and 0.75.

Once we obtain δ^*, we can directly transform it to recover $\delta(k) = \frac{\exp(\delta^*(k))-1}{\exp(\delta^*(k))+1}$, and then transform $\delta(k)$ to recover $\alpha(k)$ by $\frac{\delta(k)}{\sqrt{1-\delta^2(k)}}$.

Due to the label-switching problem, we normalize the labels of regimes to obtain accurate posterior distributions as follows $\delta(1) < \ldots < \delta(H)$. To achieve this constraint, we adopt rejection sampling.

V. Application

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). The time series is presented in Appendix B. As in Albert and Chib (1993), we consider an AR(0) model with regime switching, except that we allow changes in the skewness instead of changes in the mean. Albert and Chib (1993) show that a model without lagged coefficients (i.e., AR(0) model) is sufficient to provide a useful description for the GDP dataset. Our AR(0) Markov-switching model is as follows

$$y_t = \xi + \epsilon_t, \qquad t = 1, \cdots, T \tag{39}$$

where $T =$ is the sample size; ξ denotes the mean growth rate of real GDP; y_t is the log difference of real GDP and ϵ_t follows a skew-normal distribution as

$$p(\epsilon_t | Y_{t-1}, Z_t, S_t, \theta) = \text{skew-normal}(\epsilon_t | 0, \sigma^2, \alpha(s_t)), \tag{40}$$

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 (39) into a stochastic representation as follows

$$y_t = \xi + \delta(s_t) z_t + \sqrt{1 - \delta(s_t)^2}\nu_t, \qquad t = 1, \cdots, T, \tag{41}$$

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(0) model with Markov skewness shifts for the GDP growth rates using our maximization and MCMC procedures developed in Sections III and IV.

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 ξ we choose a Normal prior with the mean 0.00 and the standard deviation 2.00. The prior for the scale parameter (σ) follows an inverse-gamma distribution, with the mean 1.00 and the standar deviation. These two priors are rather dispersed and cover a large parameter space. The prior for the transformed shape parameters ($\delta(k), k = 1, 2$) is an uniform prior on the interval $[-1.00; 1.00]$. 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 six quarters, meaning that the average probability of staying in the same regime is equal to 0.85 and a standard deviation equal to 0.10.

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(0) model. Figure 2 displays marginal posterior density estimates for parameters of the model using normal kernel density estimates. Note also that we also provide the estimated values for $\alpha(k)$ for $k = \{1, 2\}$, the original shape parameter.

Table 1. AR(0) Markov-switching model for U.S. for real GDP growth rates

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
ξ	location	N	0.00	2.00	0.2689	0.2908	0.2858	0.1784	0.4178
σ	scale	Inv-G	1.00	1.00	1.0113	1.0014	0.9209	0.8747	1.1639
$\delta(1)$	shape	U	-1.00	1.00	-0.3533	-0.3532	-0.3642	-0.5808	-0.1125
$\delta(2)$	shape	U	-1.00	1.00	0.9505	0.9482	0.9516	0.9152	0.9717
q_{11}	prob.	B	0.85	0.10	0.8347	0.8080	0.8152	0.6976	0.8980
q_{22}	prob.	B	0.85	0.10	0.9363	0.9176	0.9209	0.8714	0.9562
$\alpha(1)$	shape	-	-	-	-0.3777	-0.3941	-0.3911	-0.7134	-0.1132
$\alpha(2)$	shape	-	-	-	3.0576	3.1222	3.0969	2.2707	4.1125
$\alpha(2) - \alpha(1)$	shape	-	-	-	3.4353	3.5163	3.4824	2.5860	4.5708

Note: N stands for Normal, B Beta, U for Uniform, 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 for the normal, beta and inverse-gamma distributions, and to the hyperparameters for the uniform distributions.

The first finding that is evident is the remarkable difference in the estimated shape parameters across the two states. The first state gives a negative value to the shape parameter (-0.3533 at the mode), while its value in the second state is positive and close to one (0.9505 at the mode). The probability intervals for these parameters are tightly concentrated, which reinforce our estimates. In last rows, we convert those parameters into $\alpha(k)$'s, $k = 1, 2$, the original shape parameter from Azzalini (1985, 1986): $\alpha(1) = -0.3777$ and $\alpha(2) = 3.0576$ at the mode. For these reasons, we label "Regime 1" as the negatively (or left) 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 0.69 and 0.89, and those for q_{22} are 0.87 and 0.95, indicating that the

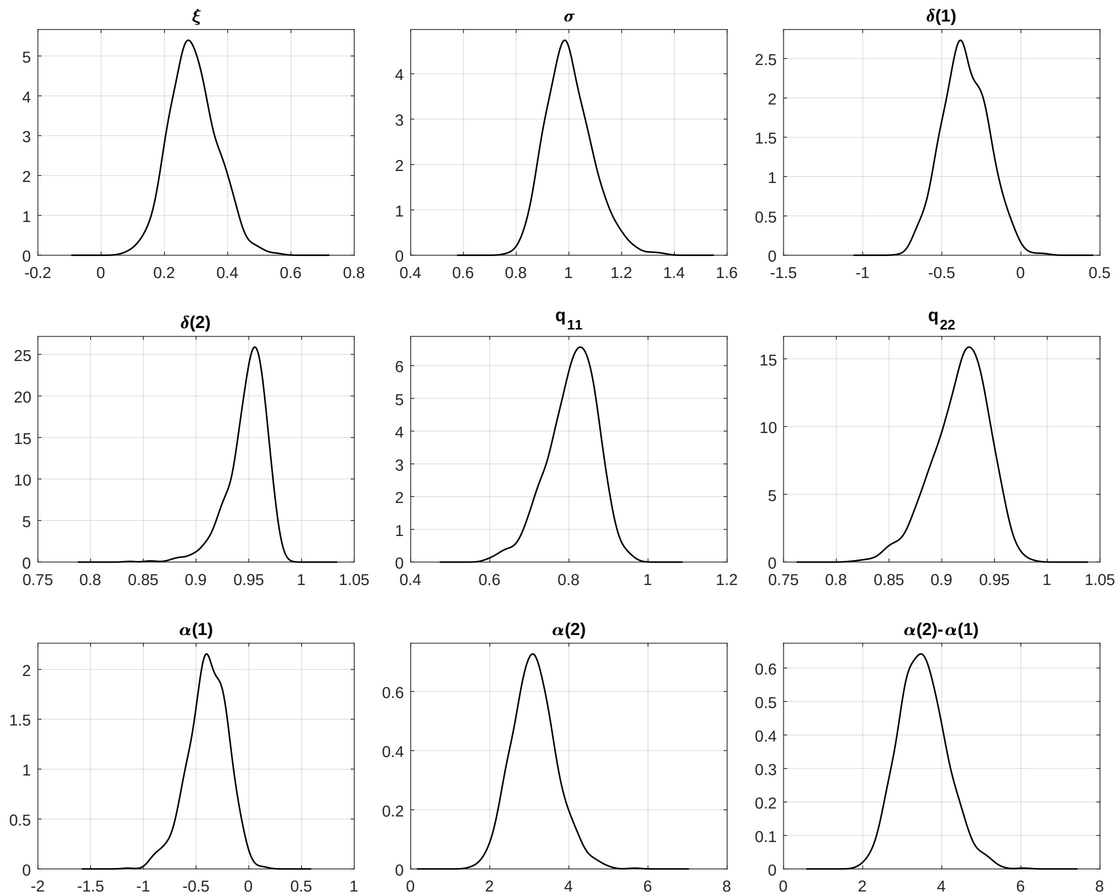

FIGURE 2. Marginal posterior densities using normal kernel density estimates.

left 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 15 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. When looking at the bottom right corner of Figure 2 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.[2] Results like this clearly imply that macroeconomic models should go beyond both linearity and gaussianity to better understand business cycle fluctuations.

Figure 3 reports the probabilities — evaluated at the mode — of being in the left skewed distribution regime over time. The probabilities are smoothed in the sense of Kim (1994); i.e., full sample information is used in getting the regime probabilities at each date. One can see from the figure that the U.S. economy has been characterized by switches between the

[2]Note that those results remain unchanged when looking at $\delta(2) - \delta(1)$ instead of $\alpha(2) - \alpha(1)$.

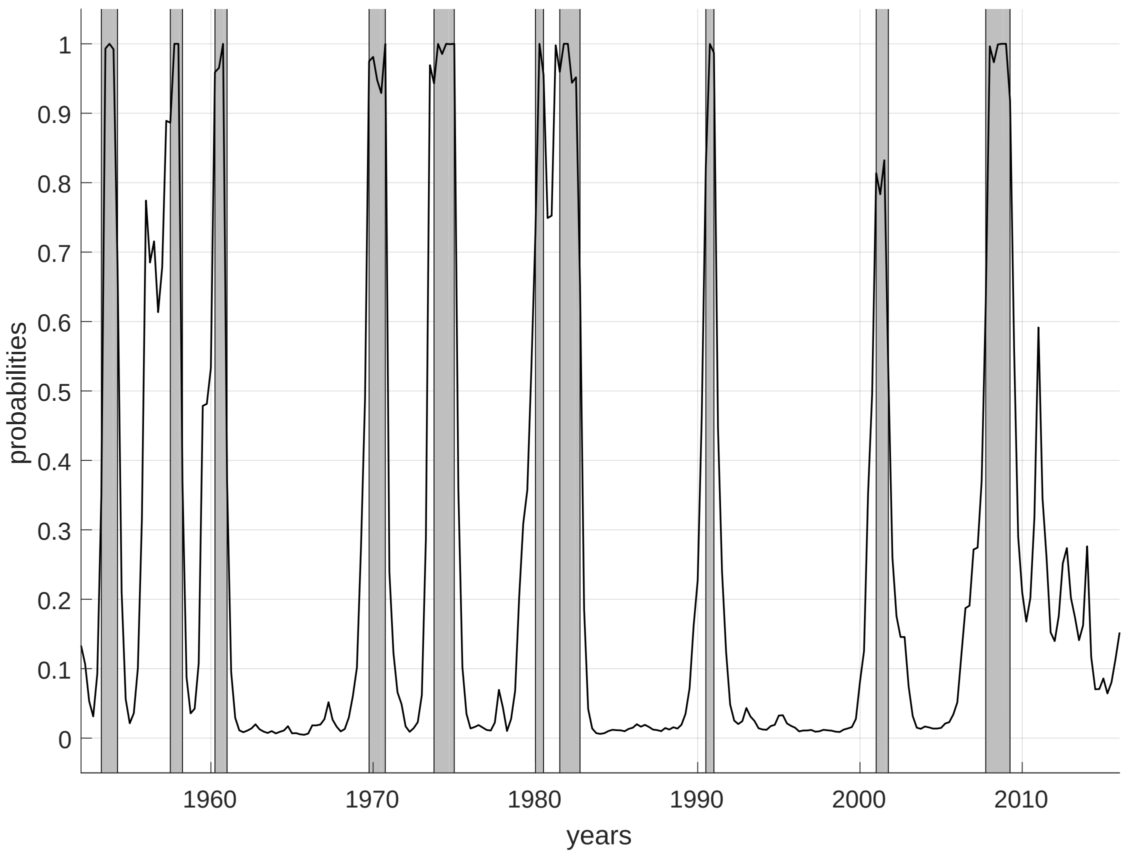

FIGURE 3. Sample period: 1952.Q1 — 2016.Q1. Smoothed probabilities of the left tail regime. The grey areas denote the NBER recessions of the United States.

left and right-tail regimes over time. In particular, the left tail 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 tends to be skewed to the left in recession phases, while the distribution is skewed to the right in expansion phases. Note also that there are two short-lived periods for which the left-skewed distribution regime does not coincide with NBER recessions: 1980.Q4-1981.Q2 and 1955.Q1-1957.Q2.

Finally, Figure 4 shows the whole distribution of GDP across regimes using our Markov-switching skew-normal model. We set the mean, the scale and the shape parameters at the mode, i.e., $\xi = 0.26$, $\sigma = 1.01$, $\delta(1) = -0.37$ and $\delta(2) = 3.05$. The dotted line is the empirical density under Regime 1 (Recession), while the solid line reports the density under Regime 2 (Expansion). One striking finding is that the variance of GDP growth is much larger during Regime 1 than during Regime 2. By cosntruction, these differences are primarly due to changes in the tail of the distribution. During business cycle expansion, the distribution of GDP growth has a lower variance and lies almost entirely within the positive region due to a large positive skewness. During recessions, instead, the distribution of GDP growth has higher variance and lies both within the negative and positive region due to a modest negative skewness. These results might seem to contradict the recent empirical finding by

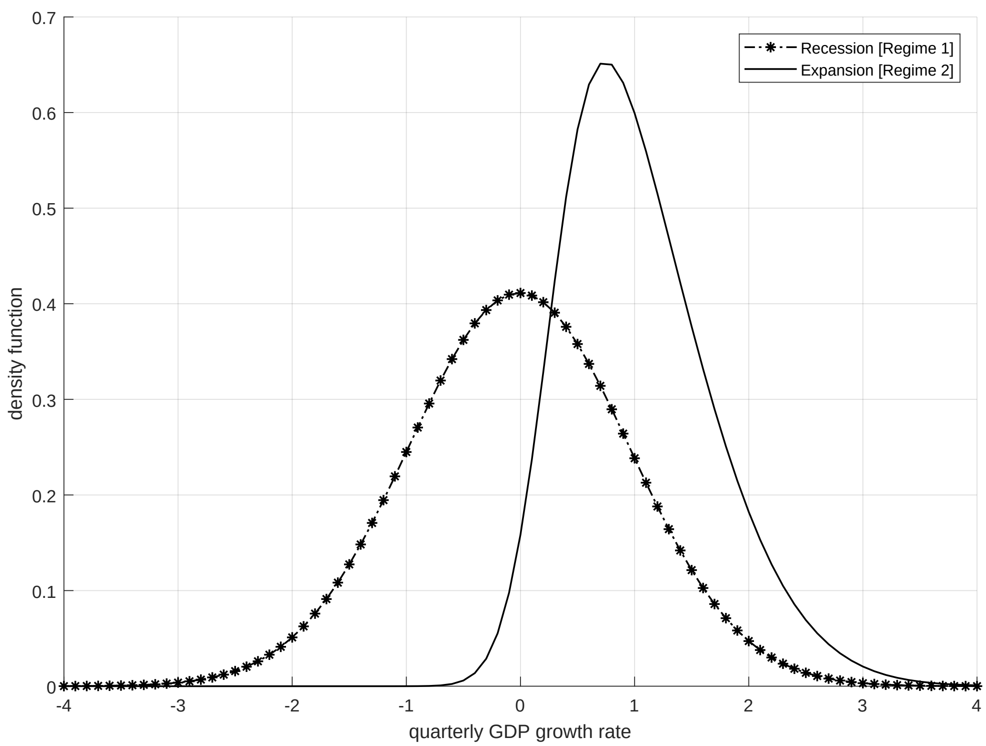

Figure 4. The distribution of quarterly GDP growth rates across regimes. It is estimated (at the mode) using the skew-normal distribution. The dotted line is the empirical density under Regime 1 (Recession), while the solid line reports the density under Regime 2 (Expansion).

Adrian, Boyarchenko, and Giannone (forthcoming) who reports a wider negative skewness during downswings, and a modest positive skewness during upswings. However, note that this study focuses on the distribution of future GDP growth as a function of current financial conditions, while here we investigate the entire (unconditional) GDP distribution.

In next section, we will show that our results remains unchanged when taking into account the possibility of regime switches in the mean of the GDP growth rates. In particular, our finding reveals that periods of low economic activity are more the result of an increase of getting a very bad outcome, rather than a decrease in the mean of output growth.

VI. Robustness

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 in the level of growth GDP. 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, with negative values in recessions and positive values in expansions.

Suppose further that the model kept the mean parameters invariant and allowed only the shape parameter to drift over time, as in our modeling framework. Since the mean 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. 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) have provided strong evidence of a structural break in the postwar U.S. real GDP growth rates in 1984.Q4. As a consequence, we think it is reasonable to model drifts independently in the skewness and the volatility over time.

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.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, we estimate our skewness switching model with a deterministic single break in the scale parameter in 1984.Q4.

VI.1. **Changes in the level of GDP growth.** We extend our benchmark model by allowing both the mean and the skewness to switch over time according to a Markov-switching process as follows

$$p(y_t|Y_{t-1}, S_t, \theta) = \text{skew-normal}(y_t|\xi(s_t), \sigma^2, \alpha(s_t)), \tag{42}$$

$$q_{i,j} = \Pr(s_t = i|s_{t-1} = j), \qquad i, j = 1, 2. \tag{43}$$

Put it differently, the times of mean changes are stochastically dependent of the times of skewness changes. Once again, we estimate the parameters of the process in equations (44) and (45) with our Bayesian methods. The MCMC procedure is slightly modified as here we also allow the mean of GDP growth 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 $\xi(s_t)$, for which we choose asymmetric priors. For the first regime, we impose a Normal prior with the mean 0.00 and the standard deviation 2.00, while for the second regime, we choose a mean equal to 1.00 and a standard deviation to 2.00. This is meant to reflect that the mean of U.S. real GDP growth rates has experienced changes over time. 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 2. AR(0) Markov-switching model for U.S. for real GDP growth rates with mean and skewness drifts

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
$\xi(1)$	location	N	0.00	2.00	0.3420	0.3536	0.3484	-0.0040	0.7176
$\xi(2)$	location	N	1.00	2.00	0.2749	0.2869	0.2847	0.1571	0.4238
σ	scale	Inv-G	1.00	1.00	1.0142	1.0213	1.0094	0.8656	1.2182
$\delta(1)$	shape	U	-1.00	1.00	-0.4496	-0.4343	-0.4420	-0.6426	-0.2069
$\delta(2)$	shape	U	-1.00	1.00	0.9523	0.9516	0.9550	0.9161	0.9771
q_{11}	prob.	B	0.85	0.10	0.8224	0.8088	0.8138	0.6895	0.9094
q_{22}	prob.	B	0.85	0.10	0.9277	0.9174	0.9220	0.8624	0.9547
$\xi(2)-\xi(1)$	location	-	-	-	-0.0671	-0.0667	-0.0486	-0.4323	0.2646
$\alpha(1)$	shape	-	-	-	-0.5034	-0.5045	-0.4927	-0.8388	-0.2115
$\alpha(2)$	shape	-	-	-	3.1213	3.2889	3.2186	2.2852	4.5921
$\alpha(2)-\alpha(1)$	shape	-	-	-	3.6247	3.7934	3.7186	2.6507	5.1442

Note: N stands for Normal, B Beta, U for Uniform, 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 for the normal, beta and inverse-gamma distributions, and to the hyperparameters for the uniform distributions.

Table 2 reports priors, modes, means, medians and 90% probability intervals for the Markov-switching AR(0) model. Results reveal that there are still important differences across regimes, where the estimate for δ is -0.44 with the probability interval $[-0.64; -0.20]$ in Regime 1, while its estimate is about 0.95 with probability interval $[0.91; 0.97]$ in Regime 2. Interestingly, it turns out that there are not any major differences in the mean of real GDP growth rates between both regimes. Indeed, the 90% error bands for $\xi(2)-\xi(1)$ lie within the negative and positive regions ($[-0.43; 0.26]$), indicating that the differences between regimes are not apparent.

Finally, although not reported[3], the estimated (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 more from changes in the skewness than in the mean.

VI.2. **Heteroskedastic variance.** We now investigate how different are our finding with a heteroskedastic variance. To do so, we allow for a single break with the onset of the Great Moderation in 1984.Q4. We consider the following model

$$p(y_t|Y_{t-1}, Z_t, S_t, \theta) = \text{skew-normal}(y_t|\xi, \sigma^2(M_t), \alpha(s_t, M_t)), \tag{44}$$

$$q_{i,j} = \Pr(s_t = i|s_{t-1} = j), \qquad i, j = 1, 2. \tag{45}$$

where

$$M_t = \begin{cases} 0 & \text{if } t < 1984.Q4 \\ 1 & \text{if otherwise} \end{cases}$$

In this setup, we assume that both the skewness ($\alpha(s_t, M_t)$) and the scale ($\sigma(M_t)$) of GDP growth rates experienced a structural break in 1984.Q4. We assume a deterministic and irreversible changes in the GDP variance, as opposed to stochastic and reversible changes. This choice is justified by the fact there is a consensus among economists on the date at which the break occured.

Estimation results are reported in Table 3. 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 regime, where its mode is equal to $\alpha(1, 1) = -0.57$ and $\alpha(1, 2) = -0.43$ in Regime 1, and $\alpha(2, 1) = 2.01$ and $\alpha(2, 2) = 2.18$ 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 $[1.57; 5.40]$ and $[1.82; 5.440]$ respectively, still indicates that 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 1.15 and 0.60, with tight probability intervals. The higher degree of volatility in the pre-1984.Q4 corroborates with Kim and Nelson (1999).

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

TABLE 3. AR(0) Markov-switching model for U.S. for real GDP growth rates with heteroskedastic variance

		Prior			Posterior				
Coefficient	Description	Density	para(1)	para(2)	Mode	Mean	Median	[5;	95]
ξ	location	N	0.00	2.00	0.4740	0.4779	0.4750	0.2637	0.7144
$\sigma(1)$	scale	Inv-G	1.00	1.00	1.1571	1.1676	1.1563	0.9968	1.3749
$\sigma(2)$	scale	Inv-G	1.00	1.00	0.6086	0.6650	0.6525	0.5662	0.7942
$\delta(1,1)$	shape	U	-1.00	1.00	−0.5001	−0.4314	−0.4385	−0.8553	0.0589
$\delta(2,1)$	shape	U	-1.00	1.00	0.8957	0.8939	0.9064	0.7770	0.9736
$\delta(1,2)$	shape	U	-1.00	1.00	−0.3987	−0.4997	−0.5064	−0.8559	−0.0993
$\delta(2,2)$	shape	U	-1.00	1.00	0.9090	0.9192	0.9262	0.8331	0.9770
q_{11}	prob.	B	0.85	0.10	0.8716	0.8367	0.8474	0.6854	0.9456
q_{22}	prob.	B	0.85	0.10	0.9145	0.8943	0.9018	0.8159	0.9480
$\alpha(1,1)$	shape	-	-	-	−0.5776	−0.6114	−0.4878	−1.6505	0.0590
$\alpha(2,1)$	shape	-	-	-	2.0146	2.3558	2.1459	1.2342	4.2643
$\alpha(1,2)$	shape	-	-	-	−0.4347	−0.7033	−0.5873	−1.6551	−0.0998
$\alpha(2,2)$	shape	-	-	-	2.1807	2.6918	2.4564	1.5062	4.5832
$\alpha(2,1)-\alpha(1,1)$	shape	-	-	-	2.5922	2.9672	2.7023	1.5768	5.4010
$\alpha(2,2)-\alpha(1,2)$	shape	-	-	-	2.6153	3.3951	3.2541	1.8561	5.4405

Note: N stands for Normal, B Beta, U for Uniform, 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 for the normal, beta and inverse-gamma distributions, and to the hyperparameters for the uniform distributions.

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. CONCLUSION

Our main goal in this paper was to provide a Gibbs sampling procedure for autoregressive time series subject to Markov skewness shifts, and use it to analyse as carefully as possible the evolution of the distribution of the growth rates of GDP in the post-war period in the United States. We have shown that the GDP distribution tends to be negatively skewed in periods of recession, meaning that the probability of getting negative growth rates is higher than the probability of observing positive growth rates. In periods of expansion, the opposite phenomenon is observed, where almost all GDP distribution lies within the positive region, due to a widening of the right tail. These asymmetric results are robust after having taken into account changes in the mean and the variance of the series.

Extending univariate dynamic models with Markov skewness shifts to a dynamic multivariate framework, like vector autoregression (VAR), 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. The works by Kim, Piger, and Startz (2008) and Chang, Choi, and Park (2017) on endogenous Markov-switching models represent the first attempt 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., N. Boyarchenko, and D. Giannone (forthcoming): "Vulnerable Growth," *American Economic Review.*

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.

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. (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. (2014): "Fluctuations in Uncertainty," *Journal of Economic Perspectives*, 28(2), 153–76.

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

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

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

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.

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.

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.

NEFTÇI, S. (1984): "Are Economic Time Series Asymmetric over the Business Cycle?," *Journal of Political Economy*, 92(2), 307–28.

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.

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

$$p(Y_T|\theta) = \prod_{t=1}^{T} p(y_t|Y_{t-1}, \theta), \tag{46}$$

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

$$p(y_t|Y_{t-1}, \theta) = \sum_{s_t \in H} p(y_t, s_t|Y_{t-1}, \theta), \tag{47}$$

$$= \sum_{s_t \in H} p(y_t|s_t, Y_{t-1}, \theta)\mathrm{Pr}[s_t|Y_{t-1}, \theta], \tag{48}$$

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

$$\Pr[s_t|Y_{t-1},\theta] = \sum_{s_{t-1}\in H} q_{s_t,s_{t-1}} \Pr(s_{t-1}|Y_{t-1},\theta), \qquad \text{for} \quad t>0, \tag{49}$$

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

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

$$\Pr[s_t|Y_t,\theta] = \frac{p(y_t, s_t|Y_{t-1},\theta)}{p(y_t|Y_{t-1},\theta)} \tag{50}$$

$$= \frac{p(y_t|s_t, Y_{t-1},\theta).\Pr(s_t|Y_{t-1},\theta)}{p(y_t|Y_{t-1},\theta)}, \qquad \text{for} \quad t>0, \tag{51}$$

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:

$$\Pr[s_t|Y_T,\theta] = \sum_{s_{t+1}\in H}^{h} \Pr[s_t, s_{t+1}|Y_T,\theta], \tag{52}$$

where

$$\Pr[s_t, s_{t+1}|Y_T,\theta] = \frac{\Pr[s_{t+1}|Y_T,\theta].\Pr[s_t|Y_T,\theta].\Pr[s_{t+1}|s_t]}{\Pr[s_{t+1}|Y_T,\theta]}. \tag{53}$$

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. Figures

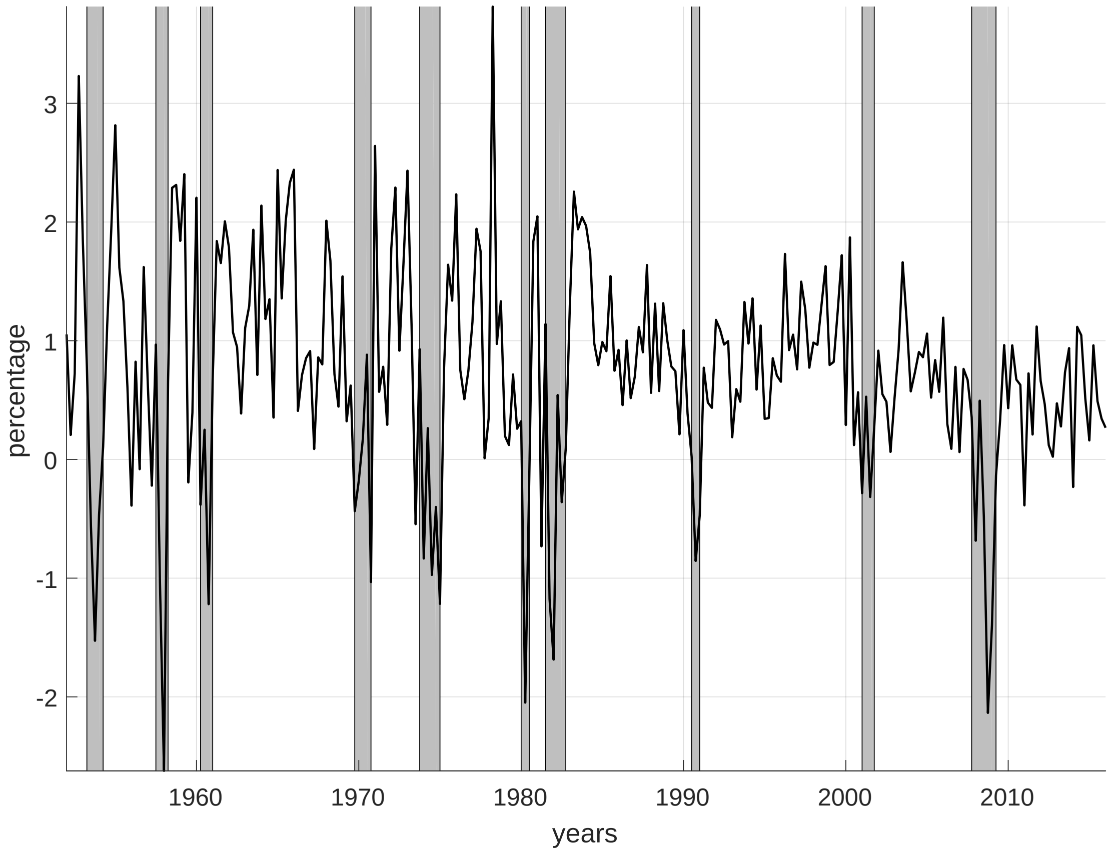

Figure 5. Sample period: 1952.Q1 — 2016.Q1. Quarterly growth rates of postwar U.S. real GDP. The grey areas denote the NBER recessions of the United States.