Estimating Growth at Risk with Skewed Stochastic Volatility Models*

Elias Wolf†

June 1, 2023

Abstract

This paper proposes a Skewed Stochastic Volatility (SSV) model to estimate asymmetric macroeconomic tail risks in the spirit of Adrian et al's (2019) seminal paper "Vulnerable Growth". In contrary to their semi-parametric approach, the SSV model captures the evolution of the conditional density of future US GDP growth in a parametric, non-linear, non-Gaussian state space model. This allows to statistically test the effect of exogenous variables on the different moments of the conditional distribution and provides a law of motion to predict future values of volatility and skewness. The model is estimated using a particle MCMC algorithm. To increase estimation accuracy, I use a tempered particle filter that takes the time-varying volatility and asymmetry of the densities into account. I find that financial conditions affect the mean, variance and skewness of the conditional distribution of future US GDP growth. With a Bayes ratio of 1612.18, the SSV model is strongly favored by the data over a symmetric Stochastic Volatility (SV) model.

Keywords: Growth at Risk, Macro-Finance, Bayesian Econometrics, Particle Filters

JEL classification: C10, E32, E58, G01.

*This paper has greatly benefited from valuable comments and suggestions by Frank Schorfheide, Frank Diebold, Carlos Montes-Galdón, Joan Paredes, Dieter Nautz, Lars Winkelmann, Helmut Lütkepohl, Yves Schüler, Lea Wolf, Max Diegel, Andrea de Polis, Laura Liu, Minchul Shin, Thorsten Drautzburg, Simon Freyaldenhoven, Matteo Cicarelli, Michelle Lenza, Mareck Jarochinski, Marta Bańbura as well as other participants of the University of Pennsylvania Econometrics Lunch Seminar, the reserach seminar at the Federal Reserve Bank of Philadelphia, the ECB Brownbag Seminar, the EABCN and Bundesbank Conference on Challenges in Empirical Macroeconomics in May 2022, the 11^{th} RCEA Money Macro Finance Conference in June 2021, and the "Topics in Time Series Econometrics"-Workshop 2020 and 2021 in Tornow. I am grateful for support from the High Performance Cluster at the Freie Universität. An earlier version of this chapter is published as FU Discussion Paper No. 2022/2.

†Contact: Freie Universität Berlin, Chair of Econometrics, Boltzmannstraße 20, 14195 Berlin, Germany. E-mail: e.wolf@fu-berlin.de

1 Introduction

John Maynard Keynes already wrote in his General Theory that economic recessions tend to be more volatile and more severe than economic expansions. Since then, different non-linear or non-Gaussian features of the business cycle have been described in the economic literature (Pitt et al., 2012). A recent study on non-linearities in US GDP growth that gained considerable attention amongst academics and policy makers alike is the seminal paper 'Vulnerable Growth' by Tobias Adrian, Nina Boryaschenko and Domenico Giannone (2019). Based on the results in the Macro-Finance literature, the authors link the non-linear and non-Gaussian features of economic growth to a country's national financial conditions. Borrowing from the Value at Risk concept in financial econometrics, they capture time-varying tail risks to GDP growth conditional on a country's national financial conditions using a semi-parametric estimation procedure. They find that a deterioration in national financial conditions increases the volatility and skewness of the conditional distribution of future GDP in the US. Most notably, their results indicate that conditional forecasting densities of US GDP growth rates are not always symmetric but left skewed in times of financial distress.

Yet, while the semi-parametric approach proposed by Adrian et al. (2019) is easy to implement, it is difficult to construct confidence intervals to statistically test the impact of national financial conditions or potentially other exogenous variables on the different moments of the conditional densities. Furthermore, since their approach does not assume a parametric law of motion, it is also not possible to conduct multi-step forecasts to predict of the evolution of time-varying risks around US-GDP growth several periods in the future. Nevertheless, as outlined in Prasad et al. (2019) the concept of Growth at Risk has provided policy makers all over world with an easy risk measure to evaluate a country's economic stability.

However, other recent studies by Hasenzagl et al. (2020) or Brownlees and Souza (2021) have put the results of Adrian et al. (2019) into question. While Hasenzagl et al. (2020) document little evidence for the effects of national financial conditions on higher moments such as variance or skewness, the results of Brownlees and Souza (2021) find that quantile regressions in the style of Adrian et al. (2019) show no predictive gains over symmetric GARCH models for macroeconomic tail risks.

Adding to the debate about time-varying asymmetries in the conditional densities of macroeconomic variables, this paper proposes a parametric modeling approach to estimate macroeconomic tail risks conditional on financial conditions. I build a Skewed Stochastic Volatility model (SSV) that can capture the variation of the full conditional density of future US GDP growth parametrically while being equally flexible. In this model, skewness arises from the assumption that errors follow a skewed normal distribution as introduced by Azzalini (2013). The model is a non-linear and non-Gaussian state space model that captures the effects of exogenous variables such as the national financial conditions index on the first and second moment of the conditional forecasting distribution of US GDP growth. Additionally, the SSV model provides a law of motion for the volatility and skewness of the conditional distributions. This makes it easy to iterate the evolution of the conditional densities forward in time to obtain multi step forecasts of tail risks. The model also nests a generic Stochastic Volatility (SV) model with symmetric densities, and does not impose skewness a priori.

Building on established Bayesian estimation methods for non-linear, non-Gaussian state space models, the SSV model can be estimated using a particle Monte Carlo Markov Chain (MCMC) algorithm that treats the model coefficients and the unobserved volatility and skewness symmetrically as random variables.[1] This allows to easily construct credible sets for both objects to conduct statistical inference. Additionally, the Bayesian estimation approach makes it possible to compare and select different model specifications based on the respective marginal data densities. To increase the efficiency of the estimation, I use the recently introduced tempered particle filter by Herbst and Schorfheide (2019) to obtain robust estimates of the likelihood and latent states. As discussed in Pitt et al. (2012) the accuracy of the particle filtering approximation is important for the efficiency of the particle MCMC algorithm. Improving on the well-documented weakness of standard bootstrap particle filters to be sensitive to extreme values (see for example Doucet et al. (2001)), the tempered particle filter is more accurate in periods of high volatility but also more computationally expensive. Building on the work of Herbst and Schorfheide (2019), I modify the tempering schedule to take the asymmetry of the measurement density into account. This results in less tempering iterations and allows to increase the targeted

[1] Flury and Shephard (2011) have documented that these estimation methods work well for SV models with symmetric densities in discrete time.

accuracy of the estimated states by reducing the runtime of the filter.

Estimating the model using US data, the SSV model provides further statistical evidence that national financial conditions have an impact on the second and third moment of the conditional forecasting distribution of future US GDP growth. Furthermore, the tempered particle filter provides significant estimates of the time variation in the variance and skewness of the forecasting density of US GDP growth rates.

The results are also in line with other recent studies in the Growth at Risk literature such as delle Monache et al. (2021) or Montes-Galdon and Ortega (2022). With a Bayes ratio of 1612.18, the SSV model is strongly favoured by the data over a symmetric SV model. Comparing the conditional densities based on entropy measures developed in Adrian et al. (2019), the higher marginal likelihood of the SSV model can be attributed to its ability to better capture the strong increase in downside risks in periods of economic turmoil. These results provide further statistical evidence that asymmetries are an essential feature of the conditional densities of macroeconomic variables in times of economic crises.

The chapter is structured as follows: Section 2 provides an overview of the Growth at Risk concept and the methodology developed by Adrian et al. (2019). The Skewed Stochastic Volatility model and its estimation is introduced in Section 3 and 4 respectively. Section 5 and 6 discuss the results for the SSV model based on US data. Section 7 concludes.

2 Growth at Risk

The concept of Growth at Risk was introduced in the seminal paper by Adrian et al. (2019) who analyse the variation of the one-period ahead forecast distribution of US GDP (gdp_{t+1}) conditional on the national financial conditions index (dubbed $nfci_t$) to analyse macro-financial risks in the US economy.[2] To obtain an estimate of the one-period ahead forecasting distribution of US-GDP, the authors develop a semi-parametric approach that consists of two steps.

In the first step, the 5, 25, 75 and 95% quantiles of the conditional distributions of gdp_{t+1} are estimated by running quantile regressions of the form

$$gdp_{t+1} = \beta_{0,\tau} + \beta_{1,\tau} nfci_t + \varepsilon_{t+1} \tag{1}$$

[2]The National Financial Conditions Index is given by the fist principal component of a large number of financial variables and released by the Federal Reserve Bank of Chicago.

where the vector of coefficients $\boldsymbol{\beta}_\tau = (\beta_{0,\tau}, \beta_{1,\tau})'$ depends on a predetermined quantile $\tau \in (0,1)$. The coefficients are found by minimizing the Koenker Bassett loss (Koenker and Bassett, 1978) defined as

$$\mathcal{L}_{KB}(\boldsymbol{\beta}_\tau) = \sum \left(\tau \cdot \mathbf{1}_{(gdp_{t+1} \geq x_t'\boldsymbol{\beta}_\tau)} |gdp_{t+1} - x_t'\boldsymbol{\beta}_\tau| + (1-\tau) \cdot \mathbf{1}_{(gdp_{t+1} < x_t'\boldsymbol{\beta}_\tau)} |gdp_{t+1} - x_t'\boldsymbol{\beta}_\tau|\right)$$

where the vector $x_t = (1, nfci_t)'$ contains the explanatory variable plus intercept and $\mathbf{1}_A(...)$ denotes the indicator function. In a second step, the authors match the predicted 5, 25, 75 and 95 percent quantiles from the regressions to the theoretical moments of the skew T distribution developed in Azzalini (2013). The respective density function is defined as

$$skew\ T(y|\xi, \omega, \alpha, \nu) = \frac{2}{\omega} \cdot t(z|\nu) \cdot T(\alpha z|\nu + 1) \quad \text{with} \quad z = \frac{y - \xi}{\omega} \tag{2}$$

with parameters ξ, ω, α, ν controlling the location, scale, shape and kurtosis of the distribution. The resulting densities are flexible and not constrained to be symmetric. Estimating the model based on US data from the 1970s up to 2017, Adrian et al. (2019) document the following properties of the one period ahead forecast distributions:

(1) Lower quantiles of the conditional forecast distribution vary a lot over time while the upper quantiles remain relatively stable.

(2) A deterioration of national financial conditions coincides with increases in the interquartile range and decreases the mean.

(3) Distributions are more symmetric in normal times and become left skewed in recessionary periods.

Based on their work, the proposed two-step approach has also been applied to analyze time-varying forecast distributions of European growth rates (de Santis and van der Veken, 2020) or other macroeconomic variables such as inflation (López-Salido and Loria, 2020). However, while the semi-parametric approach by Adrian et al. (2019) provides coefficients of the quantile regressions in the first step, it does does not provide a parametric law of motion that describes the evolution of the time variation in the volatility and asymmetry of the conditional forecast distributions. This prevents to directly capture the effect of

the national financial conditions or any other explanatory variable a researcher might include on the second and third moment. Consequently, it is also not possible to conduct statistical inference to determine the statistical relevancy of different variables on the moments of the distributions. The same is true for the estimation uncertainty of the time-varying parameters of the skew T distribution at each point in time. Furthermore, without a parametric law of motion that governs the variation in the higher moments over time it is not possible to obtain multistep forecasts of the evolution of the conditional densities and predict future risks. [3] Last but not least, it is well known that standard quantile regressions based on the Koenker Bassett Loss do not insure monotonicity in the estimated coefficients of the different quantile regressions. Hence, the regression lines for different quantiles can intersect, a problem that is commonly denoted as quantile crossing (see for example Chernozhukov et al. (2010)). To remedy these shortcomings, this paper proposes to estimate the evolution of the full forecasting density in a fully parametric way using a Skewed Stochastic Volatility Model.

3 Skewed Stochastic Volatility Model

The Skewed Stochastic Volatility Model (SSV) model is a non-linear, non-Gaussian state space model with measurement equation

$$y_t = \gamma_0 + \sum_{l=1}^{L} \gamma_l x_{t,l} + \sum_{p=1}^{P} \beta_p y_{t-p} + \varepsilon_t \quad \text{with} \quad \varepsilon_t \sim skew\, \mathcal{N}(0, \sigma_t, \alpha_t) \tag{3}$$

and latent states

$$\ln(\sigma_t) = \delta_{1,0} + \sum_{j=1}^{J_\sigma} \delta_{1,j} x_{t,j} + \sum_{k=1}^{K_\sigma} \beta_{1,k} \ln(\sigma_{t-k}) + \nu_{1,t} \tag{4}$$

$$\alpha_t = \delta_{2,0} + \sum_{j=1}^{J_\alpha} \delta_{2,j} x_{t,j} + \sum_{k=1}^{K_\alpha} \beta_{2,k} \alpha_{t-k} + \nu_{2,t} \tag{5}$$

[3] Compared to the approach in this paper, an alternative way to obtain multistep forecasts using the semi-parametric approach of Adrian et al. (2019) could be to estimate a separate model to obtain a direct forecast for each horizon in the spirit of a local projections approach. However, this method fails to capture the uncertainty around the predictions.

The innovations of the latent states $\nu_{1,t}$ and $\nu_{2,t}$ are assumed to be uncorrelated Gaussian White Noise

$$\begin{bmatrix} \nu_{1,t} \\ \nu_{2,t} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\nu_1} & 0 \\ 0 & \sigma_{\nu_2} \end{bmatrix} \right) \tag{6}$$

Most importantly, the errors in the measurement equation (3) are distributed according to the skew normal distribution of Azzalini (2013). Similar to the normal distribution it has parameters for the location (μ) and scale (σ) plus an additional shape parameter α that determines the symmetry of the density function. The probability density function of the skew normal distribution is given by

$$skew\,\mathcal{N}(y|\mu,\sigma,\alpha) = \frac{2}{\sqrt{(2\pi)}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \int_{-\infty}^{\alpha \frac{y-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}\, \mathrm{d}z \quad \text{with} \quad z = \frac{y-\mu}{\sigma}. \tag{7}$$

Figure 1 shows how different values for α affect the skewness of the distribution function. While $\alpha < 0$ results in a left skewed distribution, $\alpha > 0$ tilts the distribution to the right. Setting $\alpha = 0$ recovers the symmetric Normal distribution. The evolution of the density

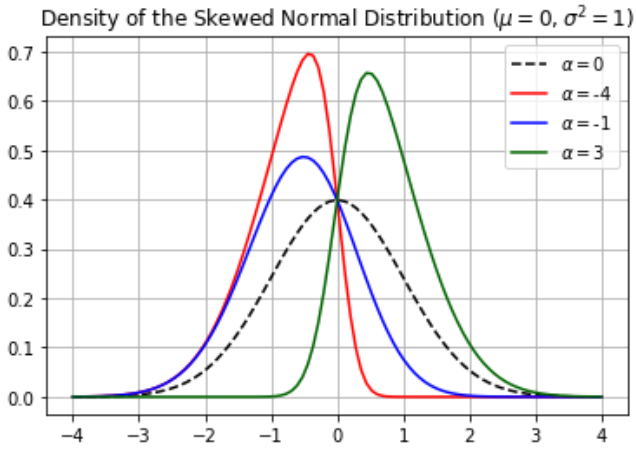

Figure 1: Skewed Normal Distribution for Different Values of the Shape Parameter α.
Notes: Negative values of α tilt the distribution to the left while positive values skew the distribution to the right. Setting $\alpha = 0$ recovers the standard normal distribution.

of the measurement errors over time is governed by state equations (4) and (5). The set of exogenous driving variables is not limited to be equal across Equations (3) - (5). Additionally, the lagged states can capture serial correlation in the evolution of the state variables. To guarantee the stability of the model I restrict the autoregressive coefficients

such that

$$\sum_{p=1}^{P} \beta_p < 1 \quad \text{and} \quad \sum_{k=1}^{K_\alpha} \beta_{1,k} < 1 \quad \text{and} \quad \sum_{k=1}^{K_\sigma} \beta_{2,k} < 1.$$

Depending on the value of the static coefficients of the model, Equation (4) will affect the interquartile range of the conditional densities while Equation (5) introduces time-varying asymmetries.

The skew normal distribution can also capture excess kurtosis in the conditional densities of the SSV model. Given the derivations of Azzalini (2013) the kurtosis of the skew normal is a function of the scale and shape parameters. Hence, skew normal distribution can capture all characteristics found by Adrian et al. (2019) while maintaining a parsimonious modelling approach and without imposing skewness a priori. Furthermore, the SSV model nests a symmetric SV model with symmetric densities if $\delta_{2,0} = \delta_{2,j} = \beta_{2,k} = \sigma_{\nu_2} = 0 \quad \forall j, k$. Last but not least, given the estimation approach outlined in Section 4, it is also possible to allow for non-zero correlation of the error terms ν_1 and ν_1 by including their covariance in the set of static parameters that need to be estimated.

Estimating the model using Bayesian methods yields credible sets for both, the static parameters of the model as well as the value of the latent states at each time t. This allows to directly test the significance of the static and time-varying parameters. Additionally, it is easy to iterate the latent states forward in time to forecast the evolution of the conditional density of y_t several periods ahead.

The results of Adrian et al. (2019) quickly spawned a number of papers that are related to the research of this paper and that can be categorized in their parameter versus observation-driven modelling approach. As defined in Koopman et al. (2016), the observation-driven framework captures time-variation of parameters as deterministic functions of lagged dependent as well as exogenous variables. Prominent examples are the ARCH and GARCH models by Engle (1982) and Bollerslev (1986). In contrast to these models, the parameter-driven approach models time-varying parameters as independent stochastic processes with idiosyncratic errors. A notable example is the stochastic volatility model or the unobserved component models discussed in Durbin and Koopman (2012). Using a parameter driven approach, Carriero et al. (2020) work with a time-varying volatility specification in large BVAR that can capture variation in the second moment of the

conditional forecasting densities. Yet, skewness can only arise in the unconditional density of GDP growth. Montes-Galdon and Ortega (2022) analyze asymmetric macroeconomic risks in the Euro Area using an SVAR model with structural errors that follow a multivariate skew normal distribution with time varying shape parameters. However, the authors limit the effect of national financial conditions to the shape parameter of the structural errors and restrict the scale parameter to be constant over time. Similarly, Iseringhausen (2021) develops a panel model with time-varying skewness. Yet, national financial conditions are restricted to affect the skewness of the conditional distributions while volatility evolves as a random walk. Eventually, Hasenzagl et al. (2020) also estimate a time-varying parameter model using Hamiltonian Monte Carlo methods and a skew T distribution. The authors find that the time-varying moments cannot be estimated precisely and are generally insignificant. However, compared to the analysis in this paper, the authors focus on marginal forecasting gains of financial variables over measures of real activity.

In the observation driven framework, Adrian et al. (2019) already include a simple GARCH-type model that can capture time variation in the first two moments. Yet, resulting conditional forecast distributions remain constrained to be symmetric and Gaussian similar to Carriero et al. (2020). Using the Generalized Autoregressive Score (GAS) model framework developed by Creal et al. (2013), the paper by delle Monache et al. (2021) also allows for time varying skewness. Furthermore, older observation driven models that seek to model time-varying skewness are proposed in the work of Hansen (1994) or Engle and Manganelli (2004).

Compared to the other studies, an additional contribution of this paper is the Bayesian estimation approach of the model using advanced particle filtering methods recently introduced by Herbst and Schorfheide (2019). In contrary to the aforementioned studies, the SSV model and the Bayesian estimation strategy can capture all features documented by Adrian et al. (2019) while remaining flexible with regards to other distribution families and more complex model specifications.

The SSV model in this paper is a parameter driven model and estimated using a combination of particle filters and Markow Chain Monte Carlo methods. To the best of my knowledge, this is the first paper that estimates a SV model that also allows for time-varying skewness using state of the art particle filtering techniques in combination

with MCMC methods. The estimation methods for the non-linear state space model are introduced in detail in the next section.

4 How to Estimate the Skewed Stochastic Volatility Model

Due to the skew normal distribution of the measurement errors as well as the non-linearity in the state equation of the scale parameter, the SSV model is a non-linear, non-Gaussian state space model. Therefore, one cannot estimate the states and static model parameters with methods such as Kalman filtering and the EM-algorithm. Yet, non-linear state space models can feasibly be estimated in a Bayesian framework using a combination of particle filters and MCMC-Algorithms (Schön et al., 2015). In particular, Kim et al. (1998) show how symmetric SV models can be estimated using a particle Metropolis Hastings algorithm. Furthermore, convergence results have been derived in Andrieu et al. (2010).

4.1 The Particle Metropolis Hastings Algorithm

In general a non-linear, non-Gaussian state space system consists of measurements y_t and latent states s_t that evolve according to the densities

$$y_t \sim p(y_t|s_t, \theta) \tag{8}$$

$$s_t \sim p(s_t|s_{t-1}, \theta). \tag{9}$$

In the SSV model, the latent states are the time-varying second and third moment of the measurement density determined by $s_t = (\ln \sigma_t, \alpha_t)$. The static model parameters are given in the vector

$$\begin{aligned}\theta = (&\gamma_0, \gamma_1, ..., \gamma_L, \beta_1, ..., \beta_P, \\ &\delta_{1,0}, \delta_{1,1}, ..., \delta_{1,J_\sigma}, \beta_{1,1}, ..., \beta_{1,K_\sigma}, \sigma_{\nu,1}, \\ &\delta_{2,0}, \delta_{2,1}, ...\delta_{2,K_\alpha}, \beta_{2,1}, ..., \beta_{2,K_\alpha}, \sigma_{\nu,2})\end{aligned}$$

and determine the dynamics of the latent states and the mean equation of the observed measurements. Estimating the model using the the particle Metropolis Hastings Algorithm consists of two steps:

Step 1: Posterior distributions of the static model parameters θ are obtained with a Metropolis Hastings sampler that generates draws from the posterior distribution

$$p(\theta|y_{1:T}, s_{1:T}) = \frac{p(y_{1:T}|s_{1:T}, \theta)p(s_{1:T}|\theta)p(\theta)}{p(y_{1:T})}. \tag{10}$$

The MCMC step of the algorithm constructs a Markov Chain $\{\theta_n\}_{n=1}^N$ of length N with stationary distribution equal to the posterior $p(\theta|y_{1:T}, s_{1:T})$.

Step 2: The second step uses a particle filter to sequentially estimate the posterior distributions of the time-varying model parameters

$$p(s_t|y_{1:t}, \theta) = \frac{p(y_t|s_t, \theta)p(s_t|y_{1:t-1}, \theta)}{\int p(y_t|s_t, \theta)p(s_t|y_{1:t-1}, \theta)\mathrm{d}s_t} \tag{11}$$

$$\text{with } p(s_t|y_{1:t-1}, \theta) = p(s_t|s_{t-1}, \theta)p(s_{t-1}|y_{1:t-1}, \theta). \tag{12}$$

using importance sampling. Doucet et al. (2001) give a thorough treatment of importance sampling and particle filters. For each point in time t, M particles $\{s_{t,i}, W_{t,i}\}_{i=1}^M$ are drawn from a proposal density $q(s_t|y_{1:t})$ and resampled based on the importance weights

$$W_{t,i} = \frac{w_{t,i}}{\sum_{i=1}^M w_{t,i}} \quad \text{with} \quad w_{t,i} = \frac{p(s_{t,i}|y_{1:t}, \theta)}{q(s_{t,i}|y_{1:t})}. \tag{13}$$

In principle, the proposal density can be chosen freely, yet a convenient choice generates draws based on the mixture density

$$q(s_t|y_{1:t}, \theta) = \sum_{i=1}^M W_{t-1,i}p(s_{t,i}|s_{t-1,i}, \theta) \quad \text{with} \quad \sum_{i=1}^M W_{t-1,i} = 1 \tag{14}$$

Given that the process for s_t is markovian, this choice yields the standard bootstrap particle filter where particles are resampled proportional to the likelihood $w_{t,i} = p(y_t|s_{t,i}, \theta)$

of the measurement y_t given the proposed states $s_{t,i}$ as discussed in Doucet et al. (2001). At each t, particles are proposed given the stochastic process of the latent states and weighted based on how well they explain the observed measurement y_t.

Conditional on a draw of parameters θ_n, the particle filter generates an estimate of the models likelihood function $p(y_{1:T}|s_{1:T}, \theta_n)$. This estimate is given by

$$\hat{p}(y_{1:T}|s_{1:T}, \theta_n) = \prod_{t=1}^{T} \frac{1}{M} \sum_{i=1}^{M} W_{t,i}. \tag{15}$$

The particle Metropolis Hastings Algorithm iterates between estimating $\hat{p}(y_{1:T}|s_{1:T}, \theta_n)$ given a draw θ_n with the particle filter and drawing a new vector of static model parameters θ_{n+1} from the posterior distribution $p(\theta|y_{1:T}, s_{1:T})$ using $\hat{p}(y_{1:T}|s_{1:T}, \theta_n)$ for the Metropolis Hastings step. Increasing the length N of the Markov chain as well as the number of particles M improves the accuracy of the estimation but results in more computational work and longer run times. As shown by Andrieu et al. (2010), the distribution of the resulting chain $\{\theta_n\}_{n=1}^{N}$ converges to the exact posterior distribution $p(\theta|y_{1:T}, s_{1:T})$ even if the likelihood function is estimated using the particle filter. Convergence results for the particle filter can be found in Doucet et al. (2001) or Chopin (2004).

Yet, it is well known that the proposal distribution of the bootstrap particle filter given by Equation (14) is suboptimal since it ignores information about the states s_t contained in y_t (Herbst and Schorfheide, 2019).

For example, given the SSV model a large increase of US GDP growth will be more likely under particles that suggest high values in volatility as well as positive skewness in the next period. However, as particles are proposed conditional on $s_{t-1,i}$ only a few particles will imply large values for the latent states if volatility and skewness are small in $t-1$. Since most proposed particles will have a low likelihood under the model, this results in a high variance of the normalized weights $W_{t,i}$. Consequently, only a few particles are resampled which leads to a poor approximation of the filtering density $p(s_t|y_{1:t}, \theta)$ and the likelihood of the static parameters $\hat{p}(y_{1:T}|s_{1:T}, \theta_n)$. This phenomenon is commonly referred to as weight degeneracy (Pitt and Shephard, 1999). Conversely, a smaller variance of $W_{t,i}$ implies a more uniform distribution of the weights $W_{t,i}$ and will lead to a better importance sampling approximation.

As discussed in Pitt et al. (2012) even though the particle MCMC algorithm is generally

unbiased, the quality of the likelihood approximation of the particle filter is crucial for its efficiency. A low approximation accuracy results in slow mixing properties of the markow chains which yields high rejection ratios and a slow convergence to the ergodic distribution (see for example Flury and Shephard (2011)).

A common measure to gauge the accuracy of the particle approximation at time t is the inefficiency ratio

$$Ineff_t = \frac{1}{M} \sum_{i=1}^{M} \left(\frac{w_{t,i}}{\frac{1}{M} \sum_{i=1}^{M} w_{t,i}} \right)^2 \tag{16}$$

where $w_{t,i}$ are the unormalized weights.[4] A high inefficiency ratio indicates that the distribution of weights is uneven such that the approximation of the target distribution is bad, while an inefficiency ratio close to one indicates evenly distributed weights and a good approximation of $\hat{p}(s_t|y_{1:t}, \theta)$ and $\hat{p}(y_{1:T}|s_{1:T}, \theta_n)$.

The recently introduced tempered particle filter by Herbst and Schorfheide (2019) controls the inefficiency ratio by sequentially adjusting the proposal distribution in each period. This greatly improves the accuracy of the estimated states and leads to a better approximation of the likelihood function in the Metropolis Hastings step. Building on annealed importance sampling that was first proposed by Neal (2001), it is a more complex but also more accurate filtering algorithm. Given the aim of this paper to estimate tail risks, the ability to handle outliers and extreme values better then standard particle filters makes the tempered particle filter a suitable method to obtain precise estimates of the variation in the higher moments. This property of the tempered particle filter is further elaborated in the next section.

4.2 Adjusting the Tempered Particle Filter

The tempered particle filter proposed by Herbst and Schorfheide (2019) adjusts the proposal distribution to the observation y_t using an adaptive version of annealed importance sampling procedure for each t. Instead of directly reweighting the particles drawn from $p(s_t|s_{t-1,i}, \theta)$ proportional to the likelihood function $p(y_t|s_{t,i}, \theta)$, the particles are sequentially adapted to a more optimal proposal via a sequence of N_ϕ bridge distributions. These

[4] It can be shown, if one can draw particles from the optimal proposal density $p(s_t|y_t, s_{t-1})$ the weights become $w_{t,i} = \frac{1}{M}$ $\forall i$ which gives $Ineff_t = 1$. In general this distribution is not available in closed form since it requires the distributions of the measurement and states to be conjugate to each other.

bridge distributions are defined by the "tempered" likelihood function $p_0(y_t|s_{t,i},\theta)$. The tempered likelihood function has an inflated variance defined as $\sigma_{t,i}/\phi_n$ with $0 < \phi_n < 1$. Intuitively, the likelihood function is initially "flattened" to ensure that the weights of the proposed particles are equal. As described in Herbst and Schorfheide (2019), the variance of the measurement equation is then sequentially reduced to its actual level while targeting a user-defined inefficiency ratio r^*. Concurrently, the particles are adapted to a better proposal distribution using a combination of importance sampling and MCMC methods (for a detailed description of the algorithm see Herbst and Schorfheide (2019), Herbst and Schorfheide (2014) or Godsill and Clapp (2001) for a simpler outline of the basic idea). In the context of the SSV model this means that for each t, the volatility is assumed to be large and subsequently shrunk towards a level that fits the data best during the tempering steps. For each tempering step a new value for ϕ_n is chosen as

$$\phi_n = \underset{\phi}{\operatorname{argmin}} \; Ineff(\phi) - r^* = 0 \tag{17}$$

until $\phi_{N_\phi} = 1$. If the set of particles proposed based on Equation (14) satisfies $Ineff(1) \leq r^*$ no tempering is required and the filtering step is equal to classic bootstrap particle filtering. Targeting a lower r^* will result in a better approximation of the latent states, but comes at the price of more tempering steps and a longer runtime.

Following the reasoning in Herbst and Schorfheide (2019), I modify the adaptive tempering schedule such that the asymmetry of $p_n(y_t|s_{t,i},\theta)$ is also taken into account. Since the skew normal distribution converges to a symmetric normal distribution for $\alpha \to 0$, starting from a symmetric and flat distribution will result in a value for ϕ_0 that is closer to 1 reducing the number of required tempering iterations. At each step, the targeted inefficiency ratio $r*$ can be achieved with a higher value for ϕ_n. More formally, in the SSV model the unnormalized weights $w_{t,i}(\phi_0)$ are given by

$$w_{t,i}(\phi_0) = \frac{2\phi_0^{1/2}}{\sqrt{2\pi}\sigma_{t,i}} \exp\left(\frac{-\phi_0(y_t-\mu_t)^2}{2\sigma_{t,i}^2}\right) \int_{-\infty}^{\alpha_{t,i}\phi_0^{3/2}\frac{(y_t-\mu_t)}{\sigma_{t,i}}} \exp\left(\frac{-z^2}{2}\right) \mathrm{d}z. \tag{18}$$

Compared to the unnormalized weights obtained from a normal distribution, the integral in Equation (18) introduces additional variation to $w_{t,i}$ which increases the inefficiency ratio. The tempering parameter ϕ_0 shrinks the upper bound of the integral towards 0

for $\lim \phi_0 \to 0$ and brings $p_0(y_t|s_{t,i}, \theta)$ closer to a normal distribution. Using Equation (16) and taking the limit for ϕ_0 shows that the Inefficiency Ratio is decreasing in ϕ and bounded from below by

$$\lim_{\phi_0 \to 0} Ineff_t(\phi_0) = \frac{\frac{1}{M}\sum_{i=1}^{M}\left(\frac{1}{\sigma_{i,t}}\right)^2}{\left(\frac{1}{M}\sum_{i=1}^{M}\frac{1}{\sigma_{i,t}}\right)^2} > 1 \tag{19}$$

by Jensen's inequality.[5] This is a special feature of the stochastic volatility model and differs from the lower bound derived in Herbst and Schorfheide (2019) for the DSGE model case. Without the stochastic volatility component, the lower bound is given by $r^* = 1$. Hence, I define the target inefficiency ratio as

$$r^* = \frac{\frac{1}{M}\sum_{i=1}^{M}\left(\frac{1}{\sigma_{i,t}}\right)^2}{\left(\frac{1}{M}\sum_{i=1}^{M}\frac{1}{\sigma_{i,t}}\right)^2} + \Delta_r \tag{20}$$

where Δ_r is set by the researcher to determine the accuracy of the filter.
Given Expression (7) and Algorithm 2 in Herbst and Schorfheide (2019), the expression for the unnormalized weights at the n^{th} tempering step is given by

$$\tilde{w}_{t,i}(\phi_n) = \left(\frac{\phi_n}{\phi_{n-1}}\right)^{\frac{1}{2}} \exp\left(\frac{-(\phi_n - \phi_{n-1})(y_t - \mu_t)^2}{2\sigma_{t,i}^2}\right)\tilde{\Lambda}_{t,i}(\phi_n) \tag{21}$$

with

$$\tilde{\Lambda}_{t,i}(\phi_n) = \frac{\int_{-\infty}^{\alpha_{t,i}\phi_n^{2/3}\frac{(y_t-\mu_t)}{\sigma_{t,i}}} \exp\left(\frac{-z^2}{2}\right)\mathrm{d}z}{\int_{-\infty}^{\alpha_{t,i}\phi_{n-1}^{2/3}\frac{(y_t-\mu_t)}{\sigma_{t,i}}} \exp\left(\frac{-z^2}{2}\right)\mathrm{d}z} \tag{22}$$

Once again, Expressions (21) and (22) show that the weights of the skew normal distribution differ from a symmetric normal density by a factor $\tilde{\Lambda}_{t,i}(\phi_n)$ that is greater or smaller than one depending on the sign of $\alpha_{t,i}(y_t - \mu_t)$ (see Appendix A.1). This introduces additional variation into the weights and increases the inefficiency ratio for a given ϕ_n compared to the weights from a standard normal distribution. Since the limit for ϕ_n is given as

$$\lim_{\phi_n \to 0} \tilde{\Lambda}_{t,i}(\phi_n) = 1, \tag{23}$$

[5]Note that this also holds for a stochastic volatility model with symmetric densities.

additionally tempering the skewness of the likelihood brings this factor closer to one and results in weights that are more uniform and less tempering steps.

The example in Figure 2 illustrates the idea of tempering and provides a comparison

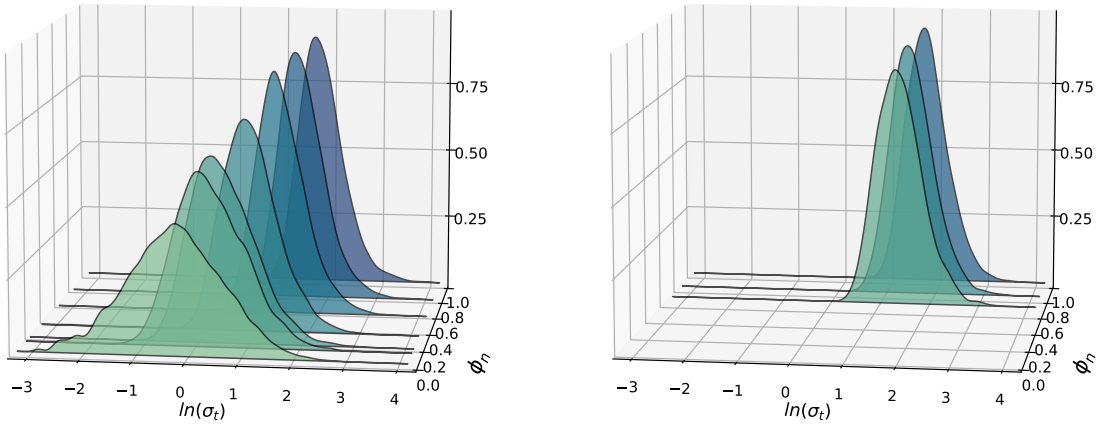

Figure 2: Tempering of the State Distributions of $\ln(\sigma_{t,i}^2)$

Notes: Illustration of tempering based on simulated data. The mean of the distribution moves from -0.2 (left) or 1.5 (right) to about 1.9. While tempering only the scale of the distribution (left) requires 7 iterations, additionally tempering the shape parameter $\alpha_{t,i}$ reduces the tempering steps to only 3 iterations (right). Furthermore, in case of skewness tempering the optimal ϕ_0 is much closer to one.

of the two different tempering schedules based on simulated data from the SSV model. The panels shows an approximation of the filtering distribution based on the proposed particles for $\ln(\sigma_{i,t}^2)$ at each tempering step. In each iteration the particles are reweighted and adjusted to the final measurement. The particles are more spread out in the beginning and slowly moved to the final filtering density at $\phi_{N_\phi} = 1$. The improvements of the skewness tempering are obvious. If the skewness of the measurement distribution is not taken into account, the mean of the filtering distribution is mutated from values of -0.2 to a final value of approximately 1.2 in 7 tempering steps. Yet, if the skewness is tempered as well it only takes 3 tempering steps and the approximation of the filtering density is more accurate at the beginning of the tempering. Furthermore, with an optimal value of 0.67, the initial ϕ_0 is already much closer to 1 in case of skewness tempering.

Figure 3 shows a comparison of the total number of tempering steps for US data from 1973 to 1983 for the tempered particle filter with and without skewness tempering. The filter is run using the model introduced in Section 3 and the estimated parameters from Section 5. As highlighted in the upper panel, this decade represents a particularly volatile period

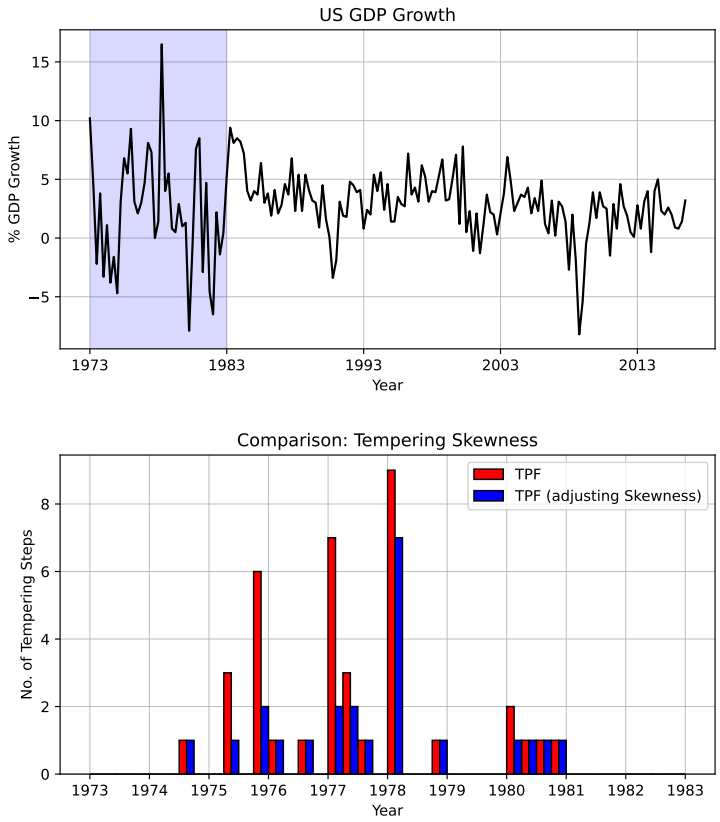

Figure 3: Total Number of Tempering Steps for both Tempering Variants
Notes: Tempering steps increase during times of high volatility. The plot shows that additionally tempering the shape of the measurement density requires fewer tempering steps.

of the sample with large jumps in US GDP growth rates. The lower panel shows the number of tempering steps that are adaptively chosen in each time period t. Tempering increases especially if the values of subsequent observations are far apart. In periods where the standard tempering schedule requires a large number of tempering steps, additionally tempering the skewness is most effective. For example, in 1975Q2, 1975Q4 or 1977Q1 skewness tempering yields a similar reduction of tempering steps as with simulated data in Figure 2. The number of tempering steps required in 1975Q2, 1975Q4 decreases by more than 60%. While exclusively tempering the scale of the measurement density takes six or seven tempering iterations additionally tempering the symmetry of the distribution requires only two iterations.

4.3 Data and Priors

The proposed model is estimated on the same data set as used by Adrian et al. (2019). Measurements y_t are chosen to be one-period ahead realizations of GDP growth (gdp_{t+1}). To compare the model with the results of Adrian et al. (2019), I use contemporaneous realizations of the national financial conditions index $nfci_t$ as exogenous driving variable in the measurement and state equations. This introduces time variation of the variance and skewness of the conditional densities of the one-period ahead US GDP growth. The lag order of the latent states is determined using the Bayes Ratio as selection criteria. I use a mixture of uninformative and informative priors on the static parameters. Table 3 in Appendix **??** gives a comprehensive overview of the prior specification of the static parameters as well as the data.

The tempered particle filter is tuned to use $M = 10,000$ particles with a targeted inefficiency ratio $\Delta_r = 0.01$ and 2 mutation steps in each tempering iteration. Draws for the static model parameters are generated using a standard random walk proposal with four chains ran in parallel on the HPC-Cluster at the Freie Universität. To increase the efficiency of the Metropolis Hastings algorithm the constrained parameters such as the autoregressive coefficients $-1 < \beta < 1$ and the variances $\sigma_{\nu_i} > 0$ are mapped to the real line using the following transformations

$$\beta = \tanh(\psi) \in [-1, 1] \tag{24}$$

$$\sigma = \exp(\zeta) \in \mathbb{R}^+ \tag{25}$$

where ψ and ζ can be drawn from the set of real numbers $\mathbb{R}$. This allows to obtain samples from the transformed unconstrained target distribution of

$$\begin{aligned}\tilde{\theta} = (&\gamma_0, \gamma_1, ..., \gamma_L, \psi_1, ..., \psi_P, \\ &\delta_{1,0}, \delta_{1,1}, ..., \delta_{1,J_\sigma}, \psi_{1,1}, ..., \psi_{1,K_\sigma}, \zeta_{\nu,1}, \\ &\delta_{2,0}, \delta_{2,1}, ...\delta_{2,J_\alpha}, \psi_{2,1}, ..., \psi_{2,K_\alpha}, \zeta_{\nu,2}) \in \mathbb{R}^S\end{aligned}$$

were $S = 5 + L + P + J_\sigma + K_\sigma + J_\alpha + K_\alpha$. As described in Schön et al. (2015) this requires to correct the acceptance ratio for the Jacobians of the inverse functions

$$\frac{\mathrm{d}\tanh^{-1}(\psi)}{\mathrm{d}\psi} = \frac{1}{1-\psi^2} \quad \text{and} \quad \frac{\mathrm{d}\log(\zeta)}{\mathrm{d}\zeta} = \frac{1}{\zeta}$$

based on the change of variables rule. The posterior distributions of the constrained model parameters can then be recovered using Equations (24) and (25). To further improve the mixing properties of the chains an initial estimate of $Var(\tilde{\theta}) = \Omega$ is obtained based on a pre-run with 5000 draws. The proposal variance is scaled to target an acceptance ratio between 20% and 30% as suggested in Roberts and Rosenthal (2001).

5 Results

Table 1 presents the estimates for the static model parameters from the particle Metropolis Hastings algorithm described in the previous section. The estimated coefficient γ_1 gives

Model Parameter	Mean	SD	q16	q84	q05	q95
γ_0	2.285	0.398	1.898	2.672	1.623	2.94
γ_1	-0.686	0.362	-1.045	-0.335	-1.311	-0.119
$\delta_{1,0}$	0.865	0.285	0.573	1.164	0.446	1.372
$\delta_{1,1}$	0.242	0.096	0.147	0.338	0.102	0.412
$\beta_{1,2}$	0.108	0.278	-0.192	0.396	-0.375	0.522
$\delta_{2,0}$	0.218	0.221	0.006	0.429	-0.143	0.595
$\delta_{2,1}$	-0.290	0.226	-0.477	-0.103	-0.603	0.042
σ_{ν_1}	0.092	0.059	0.037	0.14	0.023	0.209
σ_{ν_2}	0.020	0.020	0.006	0.032	0.004	0.058

Table 1: Posterior Means, Standard Deviations (SD) and 68% and 90% Credible Sets
Notes: The model is estimated using $N = 20,000$ draws of the tempered particle Metropolis Hastings Algorithm. The first half of the sample is discarded as burn in. The model specification containing a lagged value of α_t in the state equation of the shape parameter is strongly rejected against a model specification without an autoregressive term based on a Bayes Ratio of about 200. The marginal data densities are estimated using the modified harmonic mean estimator of Geweke (1999))

a negative impact of $nfci_t$ of about -0.69 on the one period ahead realization of gdp_{t+1}. Furthermore, with an impact of 0.24 on the scale ($\delta_{1,1}$) and -0.29 on the shape ($\delta_{2,1}$) of the skew normal distribution, the effect of national financial conditions on the different moments of the conditional densities is in line with the stylized facts described in Adrian et al. (2019). Consequently, the estimated coefficients of the SSV model imply an inverse

relationship of the second and third moment of the conditional densities. As financial conditions deteriorate, the expected growth rate decreases while the interquartile range widens and downside risks of future GDP growth increase.

However, while the 90% credible sets of the coefficients for the effect of current national financial conditions on the mean and variance of gdp_{t+1} do not overlap the zero, this is not the case for the effect on the shape parameter. Significance of the impact of $nfci_t$ on the skewness of the conditional densities is only given by the 68% credible set of $\delta_{2,1}$. This raises the question about the importance of time-varying asymmetry of the conditional densities, a topic that will be further investigated in Section 6. Additionally, Figure 4 shows the sample approximations and prior distributions of the parameters γ_0 and γ_1 in the measurement equation as well as the parameters that capture the effect of $nfci_t$ on the shape and scale of the conditional distribution of gdp_{t+1}. All posteriors are well-behaved,

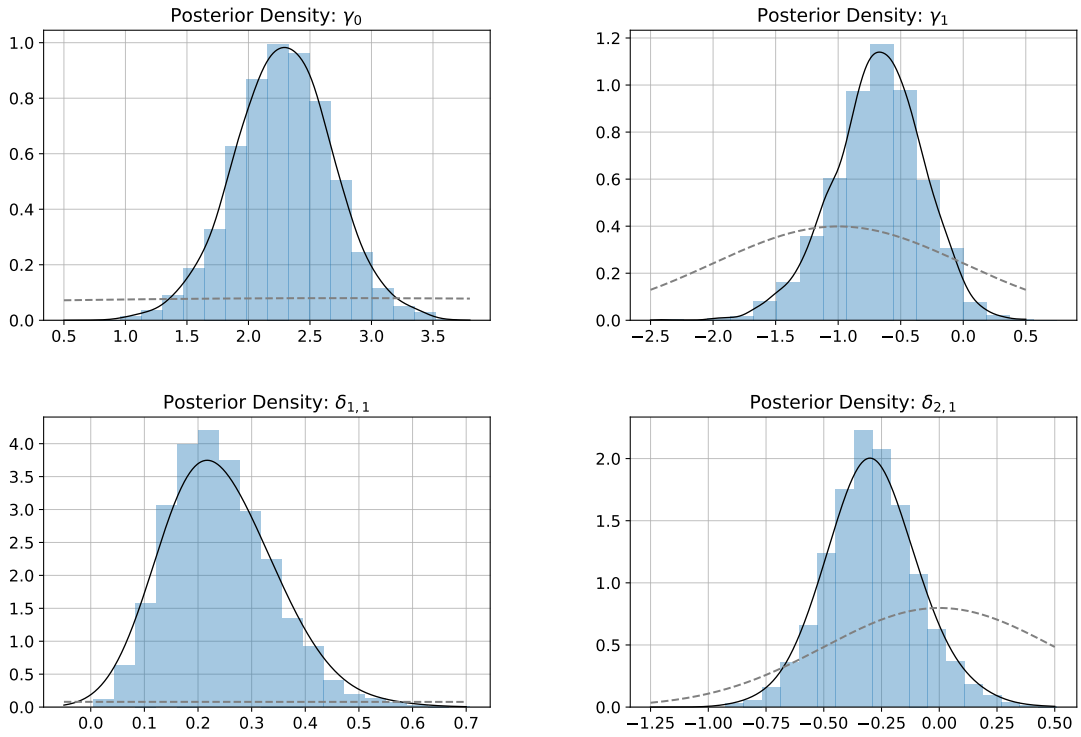

Figure 4: Posterior Distributions Obtained Using the Particle MCMC Algorithm

Notes: Posterior distributions for the parameters of the measurement equation γ_0 and γ_1 as well as the parameters that capture the effect $nfci_t$ on the scale ($\delta_{1,1}$) and shape ($\delta_{2,1}$). All posteriors are well-behaved, unimodal and clearly centered away from zero. The coefficients γ_0, γ_1 and $\delta_{1,1}$ are significantly different from zero based on the 90 % credible set constructed from the posterior draws. For $\delta_{2,1}$ the 68% credible set does not overlap the 0. Grey dashed lines indicate the prior distributions.

uni-modal and clearly centered away from zero. The distributions also differ sufficiently

from the priors indicating that the effects are well identified by the data.

Given the posterior mean estimates of the static model parameters, Figure 5 shows the filtered states and the respective 68% and 90% credible sets, an additional feature of the SSV model that is not available with the quantile approach of Adrian et al. (2019). The sharp increase in volatility and downside risks in the 1980s as well as during the financial crises in 2009 is well captured by the evolution of the two latent states. Compared to the results in Hasenzagl et al. (2020), both states indicate significant time variation in the second and third moments of the conditional densities of gdp_{t+1} based on the 90% credible sets. Yet, Hasenzagl et al. (2020) approach includes autoregressive components of GDP growth to investigate time-varying asymmetries from a forecaster's perspective who seeks to obtain additional forecasting gains by including financial variables. While the aim of this paper remains close to the work of Adrian et al. (2019), the SSV model and estimation algorithm can easily be extended to further investigate these questions and provide a further comparison with the results of Hasenzagl et al. (2020).[6] In general, the plots show that the proposed model is able to capture the stylized facts given in Section 2. However, even though the distribution is more symmetric in normal times, the estimated state of α_t also exhibits significant levels of positive skewness when levels of volatility are low. Based on the estimated states, significant upside risks to GDP growth occur during times of economic moderation for example during the late 1980s and early 1990s. These results are similar to recent results of delle Monache et al. (2021) who find evidence for cyclical behavior in the shape of the one-step ahead conditional forecast densities. Based on a trend-cycle decomposition of the latent states in their model, the authors find that conditional forecasting distributions of future US GDP growth do not only exhibit negative skewness during recessions, but become positively skewed in expansionary periods.

Eventually, Figure 6 shows the resulting conditional densities of the estimated model with the respective lower and upper 5% and 25% quantiles. The effect of the strong increase in the scale and shape parameters during the Great Recession as well as during the oil crises in the 1970s and 80s is clearly visible in the behavior of the lower quantiles. Similar to stylized fact (1) in 2, the upper quantiles of the conditional forecasting distributions remain relatively stable while the lower quantiles vary significantly over time.

[6]Including real macroeconomic variables such as for example the unemployment rate is a natural step left for further research.

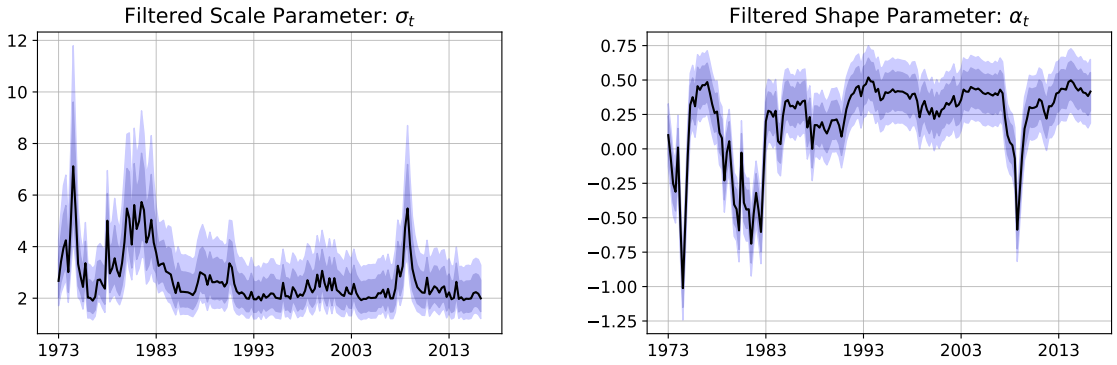

Figure 5: Filtered States Obtained with the Tempered Particle Filter

Notes: The filter is tuned to target an inefficiency ratio with $\Delta_r = 0.01$, 2 Mutation steps and $M = 10,000$ particles. The posterior means in table 1 are used for the value static model parameters. The shaded areas give the respective 68% and 90% credible sets obtained from the approximations of the filtering distributions.

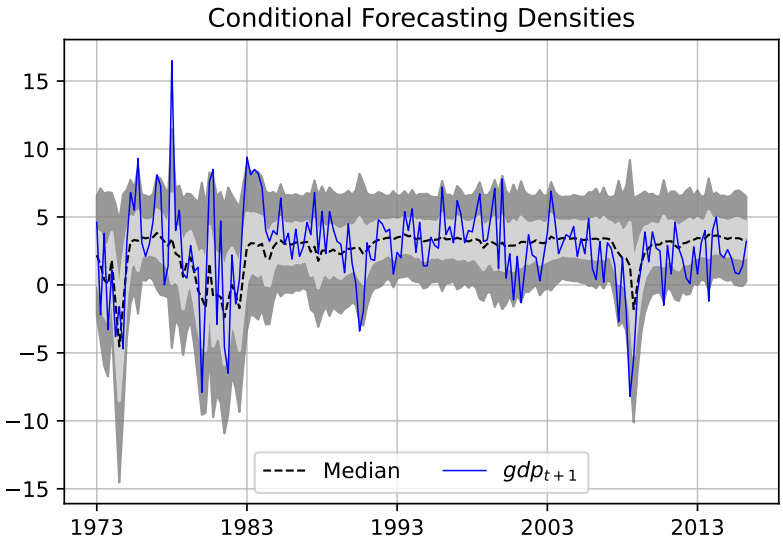

Figure 6: Conditional Distribution for One-Step Ahead GDP Growth

Notes: Lower and upper 5% and 25% percent quantiles display the same characteristics as found by Adrian et al. (2019). While the upper quantiles remain relatively stable, the lower quantiles vary strongly over time indicating increased downside risks to GDP growth in times of financial distress.

To further illustrate the difference in variation of the tail risks, Figure 7 shows the expected shortfall SF_t and expected longrise LR_t for various probability levels q. The two measures give the expected GDP growth under a specific distribution and chosen probability level q. Given the inverse CDF of the skewed Normal distribution together with the estimated parameters of the SSV model denoted by $F^{-1}_{y_{t+1}|\hat{\mu}_t,\hat{\sigma}_t,\hat{\alpha}_t}(y)$ and a chosen target probability

$q \in [0, 1]$ values for SF_t and LR_t are calculated as

$$SF_t(q) = \frac{1}{q} \int_0^q F^{-1}_{y_{t+1}|\hat{\mu}_t,\hat{\sigma}_t,\hat{\alpha}_t}(y)\mathrm{d}y \quad \text{and} \quad LR_t(q) = \frac{1}{q} \int_{1-q}^1 F^{-1}_{y_{t+1}|\hat{\mu}_t,\hat{\sigma}_t,\hat{\alpha}_t}(y)\mathrm{d}y \tag{26}$$

Similar to Value at Risk, expected Shortfall and Longrise measure tail risks under a given probability distribution. However, since they are an average over the outcomes up to a certain quantile rather than just the upper or lower bound, the behavior of the full tail is captured more comprehensively. Figure 7 shows that the complementary effects on

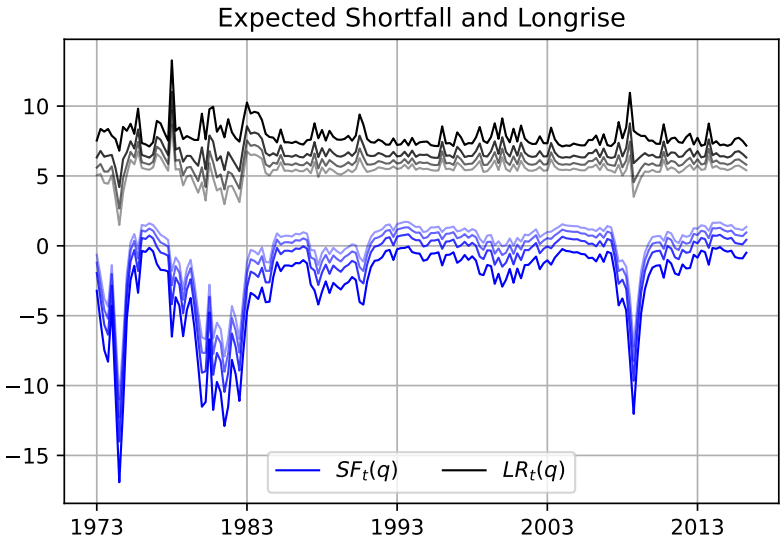

Figure 7: Expected Longrise and Shortfall

Notes: The plot shows expected shortfall/longrise for the $q = 5, 15, 25, 35$ percent quantiles. It gives the expected GDP growth in the worst/best q percent of the outcomes under the estimated skew normal distribution for each period. The plot shows that risks to the lower tails are larger in size and vary much more compared to upside risks to GDP growth.

the different moments of the distribution result in downside risks that are larger in size and vary more strongly over the full sample period. Especially the oil price shocks in the 1970s and early 1980s cause a large variation in downside risks compared to upside risks captured by the expected Longrise. During the Great Moderation the variation in the tails becomes more equal. The differences become more pronounced again during the financial crises in 2008 and 2009. Figure 7 shows that this feature is not only valid for the lower and upper 5% but also for the 15, 25 and 35% levels. However, compared to the results of Adrian et al. (2019) who find values of approximately -18% for the expected 5% shortfall during the early 1980s, the predicted tail risks of the SSV model are less severe.

6 Does Skewness Matter?

The estimation results of the SSV model yield significant coefficients for effect of national financial conditions on the the mean and the variance. However, in case of the skewness parameter the 90% credible sets for the static parameters overlap the zero. Furthermore, as argued in papers such as Adrian et al. (2020) or Carriero et al. (2020) the different behavior of the upper versus the lower quantiles can also be attributed to the inverse relationship of the mean and variance of the conditional distributions. For example, in the New Keynesian Volatiltiy Model of Adrian et al. (2020), symmetric conditional densities generate non-zero skewness in the unconditional distribution of GDP growth that closely aligns with the values observed in real world data. Additionally, based on the results of Brownlees and Souza (2021) symmetric GARCH models yield similar predictive gains as quantile regressions that allow for skewness when forecasting macroeconomic tail risks. This suggests that the inverse relationship of the mean and variance can capture much of the variation in downside risks.

To further investigate the importance of asymmetries in the conditional densities, I estimate a symmetric SV model and compare it with the SSV model based on the Bayes ratio. Since the SV model is nested in the SSV model, this corresponds to testing a joint parameter constraint on the latent state. The prior and posterior distributions of the estimated SV model are given in Appendix A.5. Since both models are estimated using Bayesian methods, the two models can easily be compared using their Bayes factor. The Bayes factor determines which model is favored by the data based on the marginal data densities $p(y|\mathcal{M}_i)$. Let $\mathcal{M}_1$ be the SSV model and $\mathcal{M}_2$ denote the symmetric SV model. The Bayes factor is given by the ratio of the two marginal data densities $p(y|\mathcal{M}_i)$

$$\frac{p(y|\mathcal{M}_1)}{p(y|\mathcal{M}_2)} = \frac{\int p(y|\theta, \mathcal{M}_1)p(\theta|\mathcal{M}_1)\mathrm{d}\theta}{\int p(y|\theta, \mathcal{M}_2)p(\theta|\mathcal{M}_2)\mathrm{d}\theta} \tag{27}$$

The Bayes factor indicates which model describes the observed data better. In general, a value larger then one indicates that $\mathcal{M}_1$ fits the data better, while values smaller than one imply a better model fit of $\mathcal{M}_2$. The marginal data densities $p(y|\mathcal{M}_i)$ are estimated

using the the modified harmonic mean estimator of Geweke (1999).[7] Based on the table of Kass and Raftery (1995) a Bayes factor lager than 10 indicates strong evidence in favor of $\mathcal{M}_1$, a Bayes factor larger than 100 implies decisive evidence. Values lower then three are considered to be inconclusive. Table 2 gives the results for the respective quantities. With a value of 1612.18 the SSV model is clearly favored by the data. Given the alterntive

Bayes Factor	log Odds	$\log p(y\|\mathcal{M}_1)$	$\log p(y\|\mathcal{M}_2)$
1612.18	7.39	-435.78	-443.16

Table 2: Model Selection and Evaluation

Notes: Bayes Factor and the log of the marginal data densities for the SSV and the SV-Model. The Bayes factor gives decisive evidence for the SSV-Model.

of a symmetric SV model, this implies that allowing for time varying skewness in the conditional densities of gdp_{t+1} increases the model fit.

Based on these results, I also compute the upside and downside entropy $\mathcal{L}_t^U$ and $\mathcal{L}_t^D$ defined in Adrian et al. (2019) to further analyze the differences of the SSV and the SV model. The upside and downside entropy compare the divergence in the probability mass of two probability distributions above and below the median. As described in Adrian et al. (2019), it is a relative measure of the divergence between two distributions in the upper and lower tails. Formally, $\mathcal{L}_t^U$ and $\mathcal{L}_t^D$ are given by

$$\mathcal{L}_{\mathcal{M}_i}^U = -\int_{-\infty}^{\hat{F}_{\mathcal{M}_i}^{-1}(0.5)} (\log \hat{g}(y) - \log \hat{f}_{\mathcal{M}_i}(y))\hat{f}_{\mathcal{M}_i}(y)\mathrm{d}y \tag{29}$$

$$\mathcal{L}_{\mathcal{M}_i}^D = -\int_{\hat{F}_{\mathcal{M}_i}^{-1}(0.5)}^{\infty} (\log \hat{g}(y) - \log \hat{f}_{\mathcal{M}_i}(y))\hat{f}_{\mathcal{M}_i}(y)\mathrm{d}y \tag{30}$$

where $\hat{g}(y)$ denotes the fitted unconditional density of GPD growth. The densities $\hat{f}_{\mathcal{M}_i}(y)$ are the conditional densities of the SSV model ($\mathcal{M}_1$) and the SV model ($\mathcal{M}_2$) given the estimated time-varying parameters. Upside entropy $\mathcal{L}_{\mathcal{M}_i}^U$ becomes positive in a given period if more probability mass is shifted to the upper tail compared the the unconditional distribution of GDP growth and vice versa. Similarly, high values for downside entropy

[7] The Harmonic Mean Estimator is given by

$$\hat{p}(y) = \frac{1}{N}\sum_{n=1}^{N} \frac{f(\theta_n)}{p(y|\theta_n)p(\theta_n)} \tag{28}$$

were $\{\theta_n\}_{n=1}^N$ are the draws obtained from the Metropolis Hastings sampler.

$\mathcal{L}^U_{\mathcal{M}_i}$ indicate that under model $\mathcal{M}_i$, there is more probability mass in the lower tail compared to the unconditional distribution. Figure 8 shows $\mathcal{L}^U_{\mathcal{M}_i}$ and $\mathcal{L}^U_{\mathcal{M}_i}$ for both models.

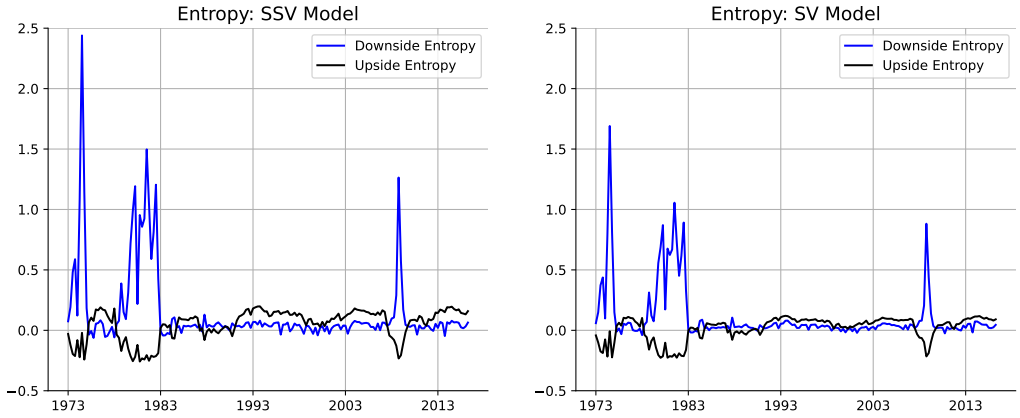

Figure 8: Upside and Downside Entropy of the SSV Model and the SV Model
Notes: Upside and Downside Entropy for the SSV model (left) and the SV model (right). Relative to the unconditional distribution, the SSV model shifts more of the probability mass to the downside compared to the SV model indicating higher tail risks.

The two plots indicate that downside entropy for the SSV model is much higher in times of economic crisis compared to SV model. Yet, the upside risks of the SSV model during the Great Moderation are visible but less pronounced compared to the downside risks during economic crises. With regards to the upside entropy, both models are fairly similar. This indicates that modelling asymmetries matters especially to appropriately capture the risks to the lower tails of the distribution. This is in line with other recent finding in the literature of asymmetric forecasting distributions. For example Montes-Galdón et al. (2022) find that introducing information on the asymmetry of forecasting densities in a BVAR strongly improves the probabilistic forecasts of GDP, inflation and core inflation in the Euro Area during times of economic crisis while there are no substantial gains in periods with stable economic conditions. In general, these results motivate future research to address the question of different risk regimes of macroeconomic variables. While models such as the SSV model can capture these characteristics, the implications of risk regimes suggest that markov switching models with exogenous driving variables that impact the transition probabilities could be another class of models to explore in subsequent studies.

7 Conclusion

This paper proposes a Skewed Stochastic Volatility (SSV) model as an alternative method to estimate Growth at Risk as introduced by Adrian et al. (2019). The SSV model can capture variation in the second and third moments of the conditional forecast distributions of US GDP growth and allows researchers to estimate and conduct statistical inference on the estimated parameters. The resulting state space model is non-linear with non-Gaussian errors and can be estimated with a Particle MCMC algorithm. I obtain accurate estimates of the model likelihood and the evolution of the latent states, using the tempered particle filter introduced by Herbst and Schorfheide (2019). Building on the adaptive tempering schedule proposed by the authors, I modify the tempering schedule to take the asymmetry of the distribution of the measurement error into account. This reduces the number tempering steps to save computational time while achieving the same accuracy. Estimating the model based on US data yields conditional forecast densities that closely resemble the findings by Adrian et al. (2019).

Exploiting the advantages of the proposed model, I find that national financial conditions have an effect on the moments of the forecasting distribution. The tempered particle filter provides significant estimates of the variation in the variance and skewness over time that imply a strong positive relationship between volatility and downside risks. Times with high volatility in growth rates coincide with an increase in risks to the lower tail of the conditional distributions. My results are also in line with results of other recent studies such as Montes-Galdon and Ortega (2022) or delle Monache et al. (2021). Compared to the findings of Hasenzagl et al. (2020), my results indicate that there is predictive content in national financial conditions for downside risks to US GDP growth and significant variation in the second and third moment of the conditional densities. I further analyze the importance of time-varying asymmetries by comparing the SSV model with a symmetric SV model using Bayesian model selection criteria. With a Bayes Factor of 1612.18, the results provide decisive evidence for the SSV model. Comparing the upside and downside entropy of the two models reveals that these advantages arise mainly from the ability of the SSV model to capture the increased tail risks in times of financial and economic turmoil.

The flexibility of the proposed SSV model and the particle MCMC algorithm allows for

further research to investigate asymmetric risks to macroeconomic variables. Given the different conclusions of Hasenzagl et al. (2020) and Adrian et al. (2020) in particular, extending the set of predictors to contain autoregressive components of GDP growth as well as other predictors can provide more insights on the predictability of time-varying asymmetries. Additionally, it is straight forward to extend the measurement equation of the SSV model to a VAR specification with additional variables and lift the exogeneity assumption of national financial conditions in the original model.

Given the active research field of macro at risk in macro-finance, empirical macroeconomics and econometrics, the SSV model provides a flexible toolkit for future research.

References

Adrian, T., N. Boyarchenko, and D. Giannone (2019). Vulnerable growth. *American Economic Review 109*(4), 1263–89.

Adrian, T., F. Duarte, N. Liang, and P. Zabczyk (2020, May). Nkv: A new keynesian model with vulnerability. *AEA Papers and Proceedings 110*, 470–76.

Andrieu, C., A. Doucet, and R. Holenstein (2010). Particle markov chain monte carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72*(3), 269–342.

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

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Brownlees, C. and A. B. Souza (2021). Backtesting global growth-at-risk. *Journal of Monetary Economics 118*, 312–330.

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

Chernozhukov, V., I. Fernández-Val, and A. Galichon (2010). Quantile and probability curves without crossing. *Econometrica 78*(3), 1093–1125.

Chopin, N. (2004). Central limit theorem for sequential monte carlo methods and its application to bayesian inference. *The Annals of Statistics 32*(6), 2385–2411.

Creal, D., S. J. Koopman, and A. Lucas (2013). Generalized Autoregressive Score Models With Applications. *Journal of Applied Econometrics 28*(5), 777–795.

de Santis, R. A. and W. van der Veken (2020). Forecasting macroeconomic risk in real time: Great and Covid-19 Recessions. Working Papers 2436, European Central Bank.

delle Monache, D., A. de Polis, and I. Petrella (2021). Modeling and forecasting macroeconomic downside risk. Working Papers 1324, Banca d'Italia.

Doucet, A., N. de Freitas, and N. J. Gordon (Eds.) (2001). *Sequential Monte Carlo Methods in Practice*. Statistics for Engineering and Information Science. Springer.

Durbin, J. and S. J. Koopman (2012, 05). *Time Series Analysis by State Space Methods*. Oxford University Press.

Engle, R. and S. Manganelli (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics 22*, 367–381.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*(4), 987–1007.

Flury, T. and N. Shephard (2011). Bayesian Inference based only on Simulated Likelihood: Particle Filter Analysis of Dynamic Economic Models. *Econometric Theory 27*(5), 933–956.

Geweke, J. (1999). Using simulation methods for bayesian econometric models: inference, development,and communication. *Econometric Reviews 18*(1), 1–73.

Godsill, S. and T. Clapp (2001). *Improvement Strategies for Monte Carlo Particle Filters*, pp. 139–158. New York, NY: Springer New York.

Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review 35*(3), 705–30.

Hasenzagl, T., M. Plagborg-Møller, R. Lucrezia, and G. Ricco (2020). When is growth at risk? Working papers, Brookings Papers on Economic Activity.

Herbst, E. and F. Schorfheide (2014). Sequential monte carlo sampling for dsge models. *Journal of Applied Econometrics 29*(7), 1073–1098.

Herbst, E. and F. Schorfheide (2019). Tempered particle filtering. *Journal of Econometrics 210*(1), 26–44.

Iseringhausen, M. (2021). A Time-varying Skewness Model for Growth-at-Risk. Working Papers 49, European Stability Mechanism.

Kass, R. E. and A. E. Raftery (1995). Bayes factors. *Journal of the American Statistical Association 90*(430), 773–795.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: Likelihood inference and comparison with arch models. *The Review of Economic Studies 65*(3), 361–393.

Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica 46*(1), 33–50.

Koopman, S. J., A. Lucas, and M. Scharth (2016, 03). Predicting Time-Varying Parameters with Parameter-Driven and Observation-Driven Models. *The Review of Economics and Statistics 98*(1), 97–110.

López-Salido, D. and F. Loria (2020, February). Inflation at Risk. Finance and Economics Discussion Series 2020-013, Board of Governors of the Federal Reserve System (U.S.).

Montes-Galdon, C. and E. Ortega (2022). Skewed SVARS: Tracking the structural sources of macroeconomic tail risks. *Advances in Econometrics (forthcoming)*.

Montes-Galdón, C., J. Paredes, and E. Wolf (2022). Conditional density forecasting: a tempered importance sampling approach. Working Paper Series 2754, European Central Bank.

Neal, R. M. (2001). Annealed importance sampling. *Statistics and Computing 11*(2), 125–139.

Pitt, M. K., R. dos Santos Silva, P. Giordani, and R. Kohn (2012). On some properties of markov chain monte carlo simulation methods based on the particle filter. *Journal of Econometrics 171*(2), 134–151.

Pitt, M. K. and N. Shephard (1999). Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association 94*(446), 590–599.

Prasad, A., S. Elekdag, P. Jeasakul, R. Lafarguette, A. Alter, A. X. Feng, C. Wang, and C. A. Gust (2019). Growth at Risk: Concept and Application in IMF Country Surveillance. Working Paper 036, International Monetary Fund.

Roberts, G. O. and J. S. Rosenthal (2001). Optimal scaling for various metropolis-hastings algorithms. *Statistical Science 16*(4), 351–367.

Schön, T. B., F. Lindsten, J. Dahlin, J. Wagberg, C. A. Naesseth, A. Svensson, and L. Dai (2015). Sequential monte carlo methods for system identification. *IFAC-PapersOnLine 48*(28), 775–786. 17th IFAC Symposium on System Identification SYSID 2015.

A Appendix

A.1 Inefficiency Ratio Under a Skew Normal Measurement Error

From Herbst and Schorfheide (2019), the weights $w_t^i(\phi_0)$ for $0 < \phi_0 \leq 1$ are given by the tempered likelihood function evaluated at the states $s_{t,i}$

Given the annealed importance sampling method described in Neal (2001) and Algorithm 2 by in Herbst and Schorfheide (2019), the expression of the unnormalized weights $w_{t,i}(\phi_n)$ are defined as the ratio of the bridge distributions

$$w_{t,i}(\phi_n) = \frac{p_n(y_t|s_t^i)}{p_{n-1}(y_t|s_t^i)} \tag{31}$$

Using expression (7) for the density of the skew normal distribution yields

$$w_{t,i}(\phi_n) = \left(\frac{\phi_n}{\phi_{n-1}}\right)^{1/2} \exp\left(\frac{-(\phi_n - \phi_{n-1})(y_t - \mu_t)^2}{2\sigma_{t,i}^2}\right) \frac{\int_{-\infty}^{\alpha_{t,i}\phi_n^{1/2}\frac{(y_t-\mu_t)}{\sigma_{t,i}}} \exp\left(\frac{-z^2}{2}\right)\mathrm{d}z}{\int_{-\infty}^{\alpha_{t,i}\phi_{n-1}^{1/2}\frac{(y_t-\mu_t)}{\sigma_{t,i}}} \exp\left(\frac{-z^2}{2}\right)\mathrm{d}z} \tag{32}$$

Expression (21) shows that in comparison to normally distributed measurement errors, the weights of the skew normal errors are scaled by a factor

$$\Lambda_{t,i}(\phi_n) = \frac{\int_{-\infty}^{\alpha_{t,i}\phi_n^{1/2}\frac{(y_t-\mu_t)}{\sigma_{t,i}}} \exp\left(\frac{-z^2}{2}\right)\mathrm{d}z}{\int_{-\infty}^{\alpha_{t,i}\phi_{n-1}^{1/2}\frac{(y_t-\mu_t)}{\sigma_{t,i}}} \exp\left(\frac{-z^2}{2}\right)\mathrm{d}z} \tag{33}$$

Additionally tempering the symmetry of the skew normal distribution modifies this factor to

$$\tilde{\Lambda}_{t,i}(\phi_n) = \frac{\int_{-\infty}^{\alpha_{t,i}\phi_n^{2/3}\frac{(y_t-\mu_t)}{\sigma_{t,i}}} \exp\left(\frac{-t^2}{2}\right)\mathrm{d}z}{\int_{-\infty}^{\alpha_{t,i}\phi_{n-1}^{2/3}\frac{(y_t-\mu_t)}{\sigma_{t,i}}} \exp\left(\frac{-z^2}{2}\right)\mathrm{d}z} \quad \begin{cases} < 1 \text{ iff } \alpha_{t,i}(y_t - \mu_t) < 0 \\ > 1 \text{ iff } \alpha_{t,i}(y_t - \mu_t) > 0 \end{cases} \quad \forall\, 0 > \phi_n > 1. \tag{34}$$

Inequality (34) holds since

$$\int_{-\infty}^{x} \exp\left(\frac{-z^2}{2}\right)\mathrm{d}z$$

is strictly monotonically increasing in x and $\phi_n > \phi_{n-1}$.

A.2 US Data

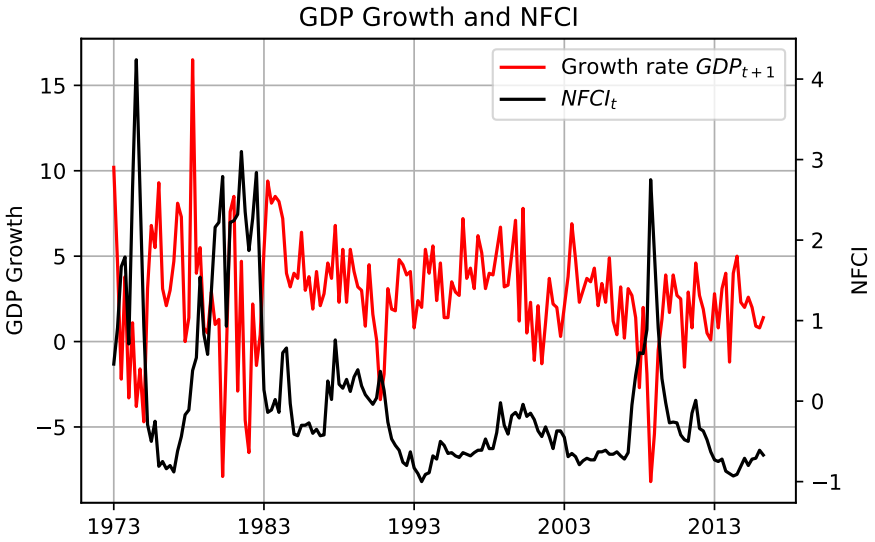

Figure 9: US GDP and National Financial Conditions Index
Notes: The sampling frequency for both series is quarterly. The sample ranges from 1973 Q1 to 2016 Q2.

A.3 Bootstrap Particle Filter vs. Tempered Particle Filter

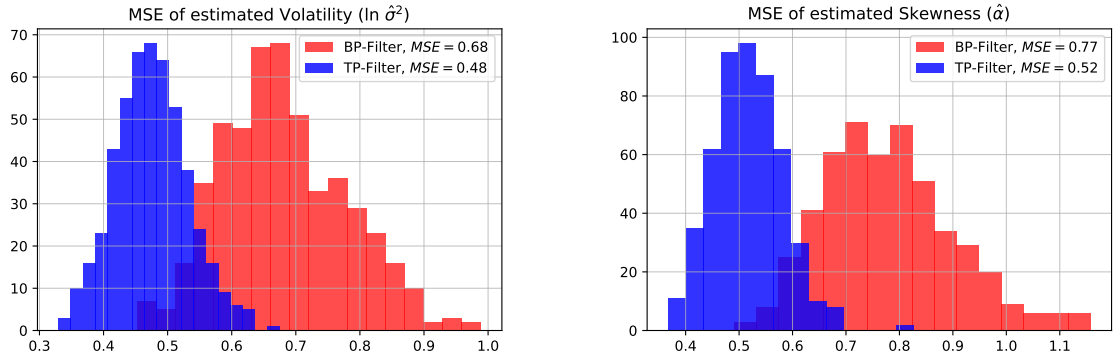

Figure 10: Mean Squared Errors of the Filtered States

Notes: Mean Squared Errors are calculated based on 500 simulations obtained with the Bootstrap Particle Filter and the Tempered Particle Filter. Tuning parameters of the Tempered Particle Filter were set to target an Inefficiency Ratio of $\Delta_r = 0.01$, 2 Mutation steps and $M = 10000$ particles. The superior performance of the Tempered Particle Filter is clear from the mean and standard errors of the distributions.

A.4 SSV-Model: Prior and Posterior Distributions

Model Parameter	Distribution	Param 1	Param 2
γ_0	N	2.69	5
γ_1	N	-1	0.5
$\delta_{1,0}$	N	0	5
$\delta_{1,1}$	N	0	5
$\beta_{1,1}$	N	0	0.5
$\delta_{2,0}$	N	0	0.5
$\delta_{2,1}$	N	0	0.5
σ_{ν_1}	IG	1	0.25
σ_{ν_2}	IG	1	0.15

Table 3: Priors for the Static Model Parameters of the SSV model.

Notes: N denotes normal priors with Param 1 and Param 2 giving mean and variances. IG denotes the inverse gamma distribution with Param 1 and Param 2 for shape α and scale β.

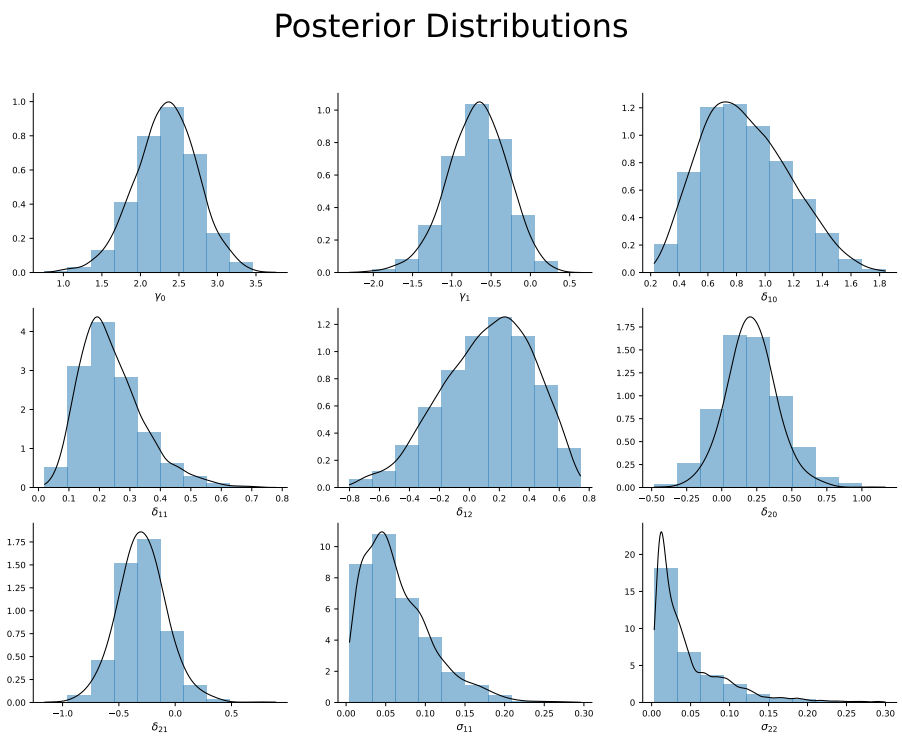

Figure 11: Skewed Stochastic Volatility Model: Posterior Distributions

Notes: Histograms and kernel density estimates of the posteriors obtained using the particle Metropolis Hastings algorithm for the SSV model.

A.5 SV-Model: Prior and Posterior Distributions

Model Parameter	Distribution	Param 1	Param 2
γ_0	N	2.69	5
γ_1	N	0	5
$\delta_{1,0}$	N	0	5
$\delta_{1,1}$	N	0	5
$\beta_{1,1}$	N	0	0.5
σ_{ν_1}	IG	1	0.25

Table 4: Priors for the Static Model Parameters of the Symmetric SV model.

Notes: N denotes normal priors with Param 1 and Param 2 giving the mean and variances. IG denotes the inverse Gamma distribution with Param 1 and Param 2 for the scale α and shape β.

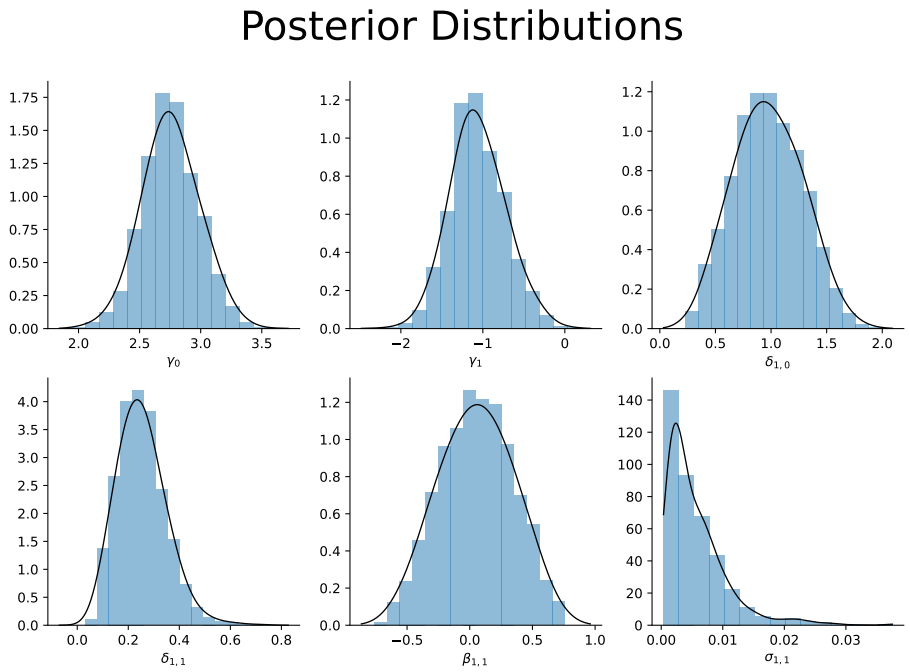

Figure 12: Stochastic Volatility Model: Posterior Distributions

Notes: Histograms and kernel density estimates of the posteriors obtained using the particle Metropolis Hastings algorithm for the symmetric SV model.