

Estimating Growth at Risk with Skewed Stochastic Volatility Models

Elias Wolf

Freie Universität Berlin

EEA-ESEM 2023

Barcelona

28/08/2023

- Asymmetric features of economic recessions and expansions have already been already described by J.M. Keynes in his General Theory.
- Since then, various non-linear and non-Gaussian features of the Business Cycle have been described in the literature (see for example Pitt et al. (2012))

- Asymmetric features of economic recessions and expansions have already been already described by J.M. Keynes in his General Theory.
- Since then, various non-linear and non-Gaussian features of the Business Cycle have been described in the literature (see for example Pitt et al. (2012))
- Vulnerable Growth by Adrian, Boyarchenko, and Giannone (2019): Relates asymmetries in future US GDP growth rates to current national financial conditions
 - Deterioration of financial conditions coincide with a decrease in future GDP growth while the volatility and skewness of the conditional distribution of future GDP Growth increase
 - Growth at Risk: Lower quantiles vary significantly over time while the upper quantiles are relatively stable

- Asymmetric features of economic recessions and expansions have already been already described by J.M. Keynes in his General Theory.
- Since then, various non-linear and non-Gaussian features of the Business Cycle have been described in the literature (see for example Pitt et al. (2012))
- Vulnerable Growth by Adrian, Boyarchenko, and Giannone (2019): Relates asymmetries in future US GDP growth rates to current national financial conditions
 - Deterioration of financial conditions coincide with a decrease in future GDP growth while the volatility and skewness of the conditional distribution of future GDP Growth increase
 - Growth at Risk: Lower quantiles vary significantly over time while the upper quantiles are relatively stable
- Semi-parametric two-step approach used by Adrian et al. (2019) makes it difficult to statistically test the empirical results
- More recent contributions of Plagborg-Møller et al. (2020) and Brownlees and Souza (2021) put the importance of time-varying moments into question.

This Paper:

- Propose a parametric Skewed Stochastic Volatility model (SSV) to estimate Growth at Risk
- Non-linear, non-Gaussian State Space Model that enables statistical inference of the estimated states and the effect of exogenous driving variables on the moments of the conditional density
- Bayesian estimation approach using a tempered Particle MCMC algorithm that allows for model comparison and selection using Bayes Ratio.
- Modification of the Tempered Particle filter of Herbst and Schorfheide (2019) to account for time-varying skewness

This Paper:

- Propose a parametric Skewed Stochastic Volatility model (SSV) to estimate Growth at Risk
- Non-linear, non-Gaussian State Space Model that enables statistical inference of the estimated states and the effect of exogenous driving variables on the moments of the conditional density
- Bayesian estimation approach using a tempered Particle MCMC algorithm that allows for model comparison and selection using Bayes Ratio.
- Modification of the Tempered Particle filter of Herbst and Schorfheide (2019) to account for time-varying skewness

Foreshadowing of results:

- Significant time variation in the second and third moment of the one-period ahead forecasting density
- SSV model chosen over symmetric SV model based on Bayes Ratio and log predictive densities
- Skewness is important to capture downside risks in times of economic turmoil

Semi-parametric approach

- **First Step:** Estimate 5, 25, 75 and 95% quantiles using Quantile Regressions:

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

where

$$\beta_{\tau} = \operatorname{argmin} \sum (\tau \cdot \mathbf{1}_{(y_t > x'_t \beta_{\tau})} |y_t - x'_t \beta_{\tau}| + (1 - \tau) \cdot \mathbf{1}_{(y_t < x'_t \beta_{\tau})} |y_t - x'_t \beta_{\tau}|)$$

- **Second Step:** Fit a skewed t-distribution to match the predicted quantiles \hat{q}

$$\hat{\mu}_t, \hat{\sigma}_t, \hat{\alpha}_t, \hat{\nu}_t = \operatorname{argmin}_{\hat{q}} \sum (\hat{q} - sT^{-1}(q | \mu_t, \sigma_t, \alpha_t, \nu_t))^2$$

Drawbacks of the semi-parametric approach:

- Time variation of the distribution is not parametrically characterized
- Difficult to conduct parameter inference or multi-step forecasts
- Potential problems of quantile crossing

This Paper: Estimate the evolution of the full forecasting density using a parametric model

Skewed Normal Distribution

PDF of the Skew Normal Distribution

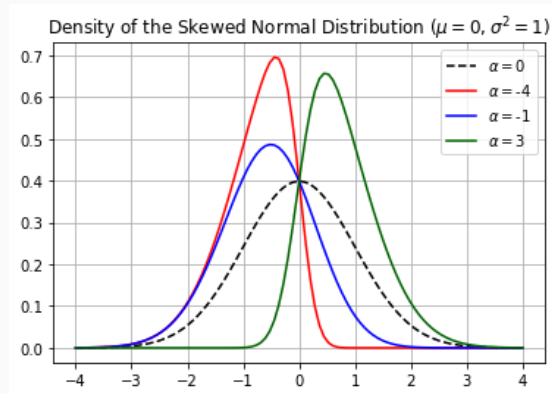
The skewed normal distribution (Azzalini (2013)):

$$f(x|\mu, \sigma, \alpha) = \frac{2}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \int_{-\infty}^{\alpha\frac{(x-\mu)}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad (1)$$

- For $\alpha = 0$ it is equal to the normal distribution.
- All three moments of interest are captured by the parameters μ, σ, α .
- Kurtosis is a non-linear combination of scale (σ) and shape (α) parameter.

→ **Parsimonious Parametrization**

→ **Estimation also works with skew T**



Skewed Normal Stochastic Volatility Model

$$y_t = \gamma_0 + \sum_{l=1}^L \gamma_l x_{t,j} + \sum_{p=1}^P \beta_p y_{t-p} + \varepsilon_t \quad \text{with} \quad \varepsilon_t \sim \text{skew } \mathcal{N}(0, \sigma_t, \alpha_t) \quad (2)$$

$$\ln(\sigma_t) = \delta_{1,0} + \sum_{j=1}^{J_\sigma} \delta_{1,j} x_{t,j} + \sum_{k=1}^{K_\sigma} \beta_{1,k} \ln(\sigma_{t-k}) + \nu_{1,t} \quad \nu_{1,t} \sim \mathcal{N}(0, \sigma_{\nu_1}) \quad (3)$$

$$\alpha_t = \delta_{2,0} + \sum_{j=1}^{J_\alpha} \delta_{2,j} x_{t,j} + \sum_{k=1}^{K_\alpha} \beta_{2,k} \alpha_{t-k} + \nu_{2,t} \quad \nu_{2,t} \sim \mathcal{N}(0, \sigma_{\nu_2}) \quad (4)$$

- National financial conditions can affect all moments of the one-period forecasting density
 - Shift in mean in (2)
 - Spread of quantile range (3)
 - Skewness of the distribution (4)
- SSV model nests a symmetric stochastic volatility (SV) model $\delta_{2,0} = \delta_{2,j} = \beta_{2,k} = \sigma_{\nu_2} = 0 \quad \forall j, k$
- Equations (2)-(4) form a non-linear, non-Gaussian state space model

Modelling time-varying skewness in conditional densities of time series:

Observation-driven models:

1. ARCH and GARCH models: Hansen (1994) Engle and Manganelli (2004) and Engle (2011)
2. Monache, Polis, and Petrella (2021): Generalized Autoregressive Score Model (Creal, Koopman, and Lucas (2013))

Parameter-driven models:

1. Iseringhausen (2021): Time varying skewness model, nfc_i_t only impacts skewness
2. Montes-Galdon and Ortega (2022): Bayesian VAR with time-varying skewness but no impact of nfc_i_t on volatility

Stochastic Volatility Model with Skewness

- Non-Gaussianity/non-linearity generally prevent use of Kalman filter and EM (Durbin and Koopman (2001))
- **Metropolis Hastings** (MH) with **Particle Filtering** (PF) for Stochastic Volatility Models (e.g. Doucet, Freitas, and Gordon (2001), Andrieu, Doucet, and Holenstein (2010), Flury and Shephard (2011))

Particle Metropolis Hastings

Static model parameters (MH):

$$\theta = (\gamma_0, \dots, \beta_1, \dots, \delta_{1,0}, \dots, \beta_{1,1}, \dots, \sigma_{\nu,1}, \delta_{2,0}, \dots, \beta_{2,1}, \dots, \sigma_{\nu,2})$$

$$\text{Posterior: } p(\theta|y_{1:T}, s_{1:T}) = \frac{p(y_{1:T}|\theta, s_{1:T})p(s_{1:T}|\theta)p(\theta)}{p(y_{1:T})}$$

Time-varying model parameters (PF):

$$s_t = (\ln \sigma_t, \alpha_t)$$

$$\text{Filtering Distribution: } p(s_t|y_{1:t}, \theta) = \frac{p(y_t|s_t, \theta)p(s_t|y_{1:t-1}, \theta)}{p(y_t|y_{1:t-1}, \theta)}$$

Estimate States: Tempered Particle Filter

- Particle Filter sequentially approximates the filtering distribution using importance sampling (15)
- Common choice of proposal distribution depends on the distributions of the states

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

which gives $W_{t,i} \propto p(y_t|s_{t,i}, \theta)$

- In case of outliers and extreme values for y_t , the proposal is not optimal and particles deteriorate.
- The quality of the particle approximation $\{s_{t,i}, W_{t,i}\}_{i=1}^M$ at time t is gauged by the **inefficiency ratio**

$$Ineff_t = \frac{1}{M} \sum_{i=1}^M W_{i,t}^2$$

- Tempered Particle Filter (Herbst and Schorfheide (2019)) uses **annealed importance sampling** to improve the approximation of the filtering distribution if the Inefficiency Ratio becomes too high.

Especially if GDP growth is volatile, the tempered particle filter becomes more accurate compared to the Bootstrap particle filter. MSE

Tempered Particle Filter I

To increase efficiency of the filter for the SSV I modify the tempering schedule to jointly temper the variance and tilt the density towards the actual level:

Intuition: Start from a **symmetric distribution** with **large variance** and temper the likelihood until

$\{s_{t,i}\}_{i=0}^M$ fits the data well. [Details](#)

- Start with initial weights:

$$W_{t,i}(\phi_0) \propto p_0(y_t | s_{t,i}, \theta) = \text{skew } \mathcal{N}(y_t | \mu_t, \sigma_{t,i} / \phi_0, \phi_0 \alpha_{t,i}) \quad \text{with} \quad 0 < \phi_0 < 1 \quad \text{and} \quad \lim_{n \rightarrow N_\phi} \phi_n = 1$$

- While $n < N_{\phi_N}$:

- Resample $\tilde{s}_{t,i} \sim \mathcal{MN}(\{s_{t,i}\}_{i=1}^M | \{\tilde{W}_{t,i}(\phi_n)\}_{i=0}^M)$ with $\tilde{W}_{t,i}(\phi_n) \propto p_n(y_t | s_{t,i})$
- Find new $\phi_{n,t}$ such that

$$\phi_{n,t} = \operatorname{argmin} \frac{1}{M} \sum_i \left[\frac{W_{i,n}(\phi_n)}{\frac{1}{M} \sum_{i=1}^M W_{i,n}(\phi_n)} \right]^2 - r^* \quad \text{where} \quad r^* = \frac{\frac{1}{M} \sum_{i=1}^M \left(\frac{1}{\sigma_{i,t}} \right)^2}{\left(\frac{1}{M} \sum_{i=1}^M \frac{1}{\sigma_{i,t}} \right)^2} + \Delta_r$$

- Mutate $\tilde{s}_{t,i} \rightarrow \hat{s}_{t,i}$ using a Transition Kernel

$$\hat{s}_{t,i} \sim K_n(\hat{s}_t | \tilde{s}_t, \hat{s}_{t-1})$$

with invariant distribution $p_n(y_t | s_{t,i}, \theta)$

- If $\phi_{n,t} = 1$ set $n = N_{\phi_N}$

Tempering the Skewness I

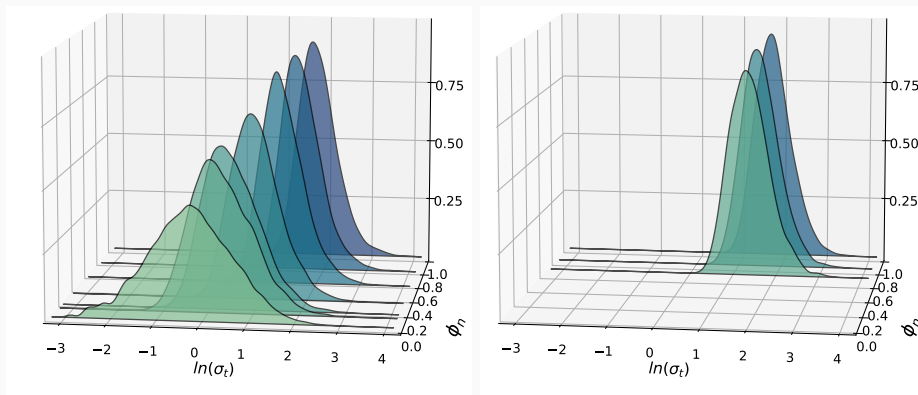


Figure 1: Particle approximation of each bridge distribution based on simulated data. Each density represents an approximation of the filtering distribution of $\log \sigma_t$. The mean of the distribution moves from -0.19 (left) or 1.46 (right) to about 1.73

- Tempering only the scale of the distribution (left) results in 7 iterations, additionally tempering the shape parameter $\alpha_{t,i}$ reduces the tempering steps to only 3 iterations (right).
- ϕ_0 is significantly smaller if only the variance is tempered (0.00049 vs. 0.67).

Tempering the Skewness: US Data

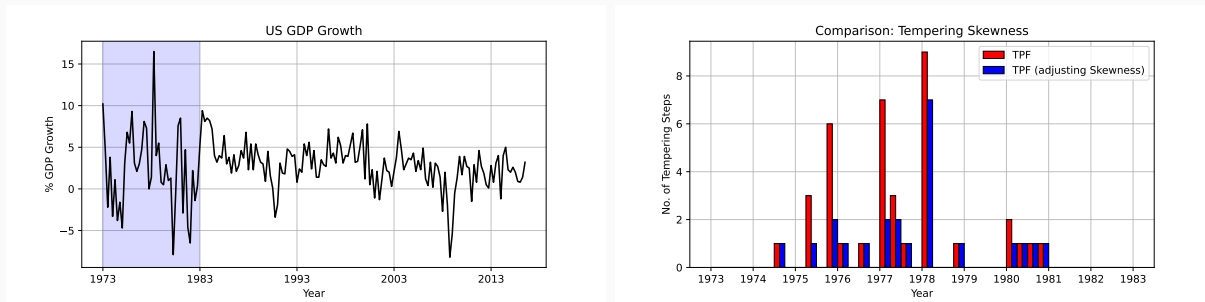


Figure 2: Additionally tempering the shape of the measurement density requires fewer tempering steps.

- Tempering steps increase during times of high volatility.
- In 1975Q2, 1975Q4 or 1977Q1, the number of tempering steps required decreases by more than 60 %
- Based on Monte Carlo study, additionally tempering the skewness decreases number of tempering steps and run time by about 25%

- Model is estimated based on data from Adrian et al. (2019) Data
- y_t is the one-period ahead US GDP growth rate gdp_{t+1}
- $nfcit_t$ is exogenous variable in equations for the level, volatility and skewness
- Lagged values in state equations are chosen based on Bayes ratio
- Mixture of diffuse and mildly informative priors on measurement equation. Priors
- To avoid discarding draws of constrained parameters, the model is reparametrized:

$$\psi_i = \tan(\beta_i) \quad \text{and} \quad \xi_i = \log(\sigma_{\nu_i})$$

to sample from the target distribution of $\tilde{\theta} \in \mathbb{R}^S$

- To improve mixing properties of the chain an estimate $\hat{\Sigma}$ is obtained from a pre-run of the algorithm.
- 4 chains with 10000 draws each are run in parallel, first half of each chain is discarded as burn-in sample.

Estimated Skewed Normal Stochastic Volatility Model

$$gdp_{t+1} = 2.29 - 0.69nfcit + \varepsilon_{t+1} \quad \text{with} \quad \varepsilon_t \sim \text{skewed } \mathcal{N}(0, \sigma_t, \alpha_t) \quad (5)$$

$$\ln(\sigma_t) = 0.87 + 0.24nfcit + 0.11 \ln(\sigma_{t-1}) + \nu_{1,t} \quad (6)$$

$$\alpha_t = 0.22 - 0.29nfcit + \nu_{2,t} \quad (7)$$

$$\nu_{1,t} \sim \mathcal{N}(0, 0.092) \quad \text{and} \quad \nu_{2,t} \sim \mathcal{N}(0, 0.020)$$

- Effect of national financial conditions on the different moments of the forecasting densities is in line with the stylized facts described in Adrian, Boyarchenko, and Giannone (2019)
- As financial conditions deteriorate, the expected growth rate decreases while the interquartile range and downside risks to GDP growth increase.

Posterior Distributions

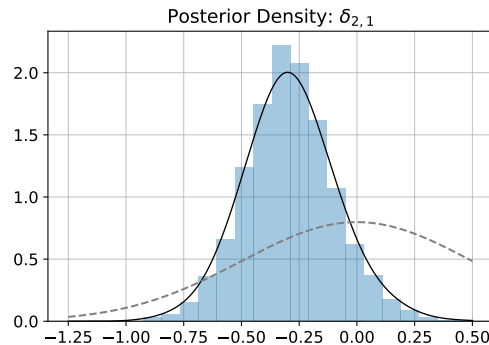
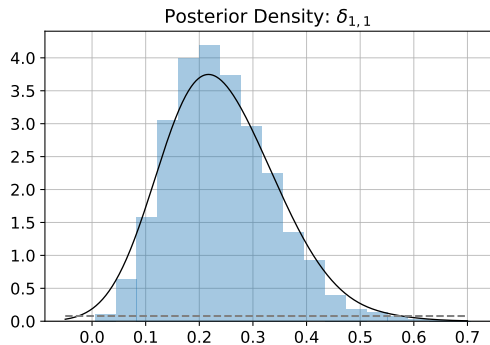


Figure 3: Posterior distributions of $\delta_{1,1}$ and $\delta_{2,1}$ obtained with the MCMC-Algorithm based on 20000 draws. Mean Equation

Equation	Param	Mean	SD	[0.16 , 0.84]	[0.05 , 0.95]
Intercept	γ_0	2.29	0.4	[1.9, 2.67]	[1.623, 2.94]
Mean Effect	γ_1	-0.69	0.36	[-1.05, -0.34]	[-1.31, -0.12]
Volatility	$\delta_{1,1}$	0.24	0.1	[0.15, 0.34]	[0.1, 0.41]
Skewness	$\delta_{2,1}$	-0.29	0.23	[-0.48, -0.1]	[-0.6, 0.04]

Time-varying Parameters

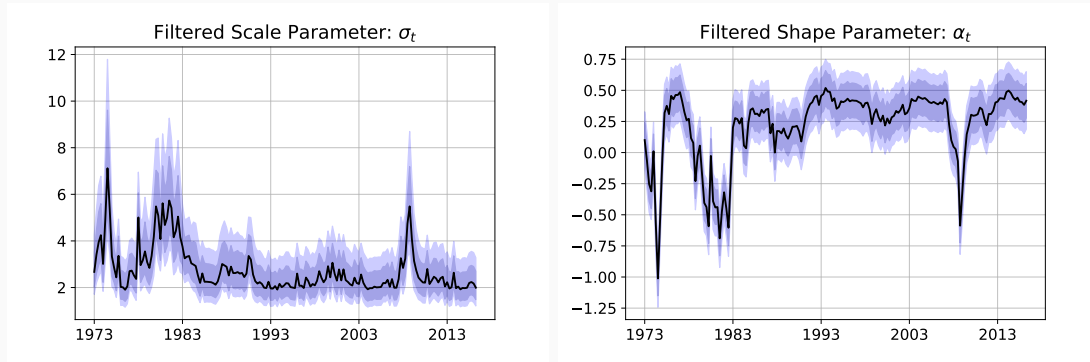


Figure 4: Filtered states given mean posterior values. The Tempered particle filter is tuned to target an Inefficiency Ratio with $\Delta_r = 0.01$, using 2 Mutation steps and $M = 10000$ particles. Shaded areas denote 68% and 90% credible sets.

- Significant variation of the second and third moment based on 68% and 90% credible sets.
- Inverse relationship \rightarrow Estimated volatility increases while skewness decreases.
- Skewness is not centered around 0 \rightarrow Upside risks in times of economic moderation (Similar to Monache, Polis, and Petrella (2021))

Growth at Risk: Conditional Densities

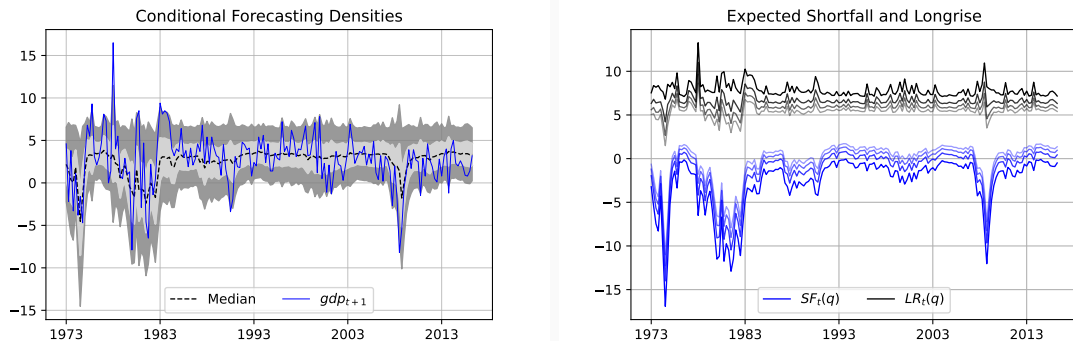


Figure 5: Right: Estimated time-varying upper and lower 5% and 25% quantiles of the conditional forecasting densities. **Left:** Expected shortfall/longrise ($SF_t(q) = \frac{1}{q} \int_0^q F_{y_{t+1}|\hat{\mu}_t, \hat{\sigma}_t, \hat{\alpha}_t}^{-1}(x) dx$ and $LR_t(q) = \frac{1}{q} \int_{1-q}^1 F_{y_{t+1}|\hat{\mu}_t, \hat{\sigma}_t, \hat{\alpha}_t}^{-1}(x) dx$) for the $q = 5, 15, 25, 35$ percent quantiles.

- Conditional forecasting densities exhibit characteristics in Adrian, Boyarchenko, and Giannone (2019) with stable upper quantiles and high variation in the lower quantiles.
- Downside risks are larger in size with a higher variance, especially in 1970s and 80s.

Does Skewness Matter?

Based on the estimated model results for asymmetric densities is still mixed

- Filtered state for $\hat{\alpha}_t$ are significantly different from 0
- 90% credible sets for the static parameters of the skewness equation overlap the zero.
- Stability of the upper quantiles can also be attributed to the inverse relationship of the mean and variance (see Adrian, Duarte, et al. (2020), Carriero, Clark, and Marcellino (2020) or Caldara, Scotti, and Zhong (2021))

Estimate a symmetric Stochastic Volatility (SV) model and compare it with the SSV model based on the Bayes Ratio and the log data densities:

- Imposing the restriction that $\delta_{2,0} = \delta_{2,1} = \sigma_{\nu_2} = 0$

Bayes Factor	log Odds	$\log p(y \mathcal{M}_{SSV})$	$\log p(y \mathcal{M}_{SV})$
1612.18	7.38	-435.78	-443.16

Table 1: 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.

Upside and Downside Entropy

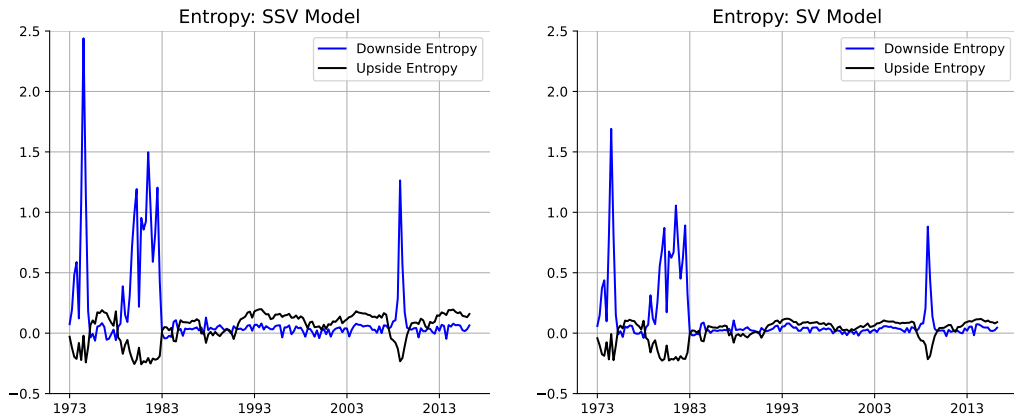


Figure 6: Upside and Downside Entropy for the SSV and SV model. Entropy

- Upside entropy is fairly equal for both models but differences in downside entropy.
- Asymmetries matter in times of economic crisis → Similar implications as in Montes-Galdón, Paredes and Wolf (2022)

Estimating a stochastic volatility model with skewness using the tempered Particle MCMC algorithm yields the following results:

- Point estimates of the parameters match the implications of other findings in literature.
- Effect of $nfcit$ on mean and volatility is significant given the 90 % -Credible Interval
- The estimated densities exhibit the features found by Adrian et al. 2019
- SSV model is favored by the data compared to SV model.
- Skewness tempering decreases run time of the tempered particle filter for asymmetric measurement densities.

Ongoing research:

- Explore larger set of exogenous variables (e.g. real economy factor) and European Data. EA-Preliminary
- Multivariate model with endogenous $nfcit$ (joint work with Montes-Galdón and Ortega).
- Joined estimation of static parameters and latent states combining Tempered Importance Sampling and Particle Learning (Carvalho et al. (2010)).

Additional Material

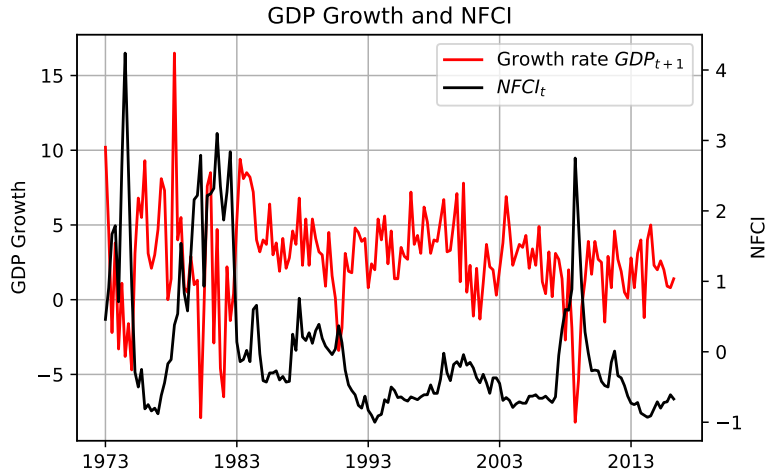


Figure 7: US GDP growth rates and NFCI index from 1973Q1 to 2016Q4.

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
$\delta_{1,2}$	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 2: Priors for the static model parameters in the Metropolis Hastings Algorithm. 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 α and β .

Importance Sampling

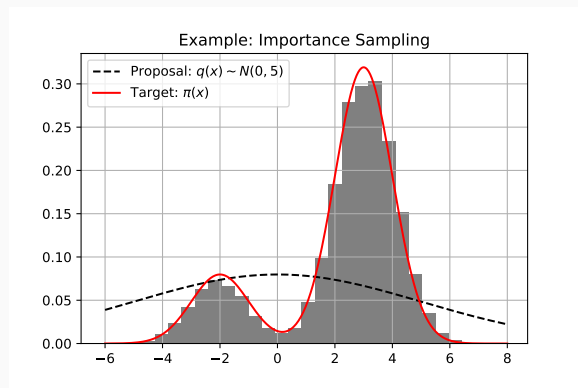
Idea: Approximate an integral over a complex target distribution $\pi(x)$ by weighted draws from a simpler proposal distribution $q(x)$.

Example: Expected Value $E_{\pi}[x]$

$$\begin{aligned} E_{\pi}[x] &= \int x_i \pi(x_i) dx \\ &= \int x_i \underbrace{\frac{\pi(x_i)}{q(x_i)}}_{w_i} q(x_i) dx \\ &\approx \frac{1}{T} \sum_{i=1}^T W_i x_i, \quad x_i \sim q(x) \end{aligned}$$

Draws from $q(x)$ are reweighted using the normalized importance weights

$$W_i = \frac{w_i}{\sum w_i}$$



Approximation of mixed Gaussian using 50000 weighted draws from a univariate normal distribution.

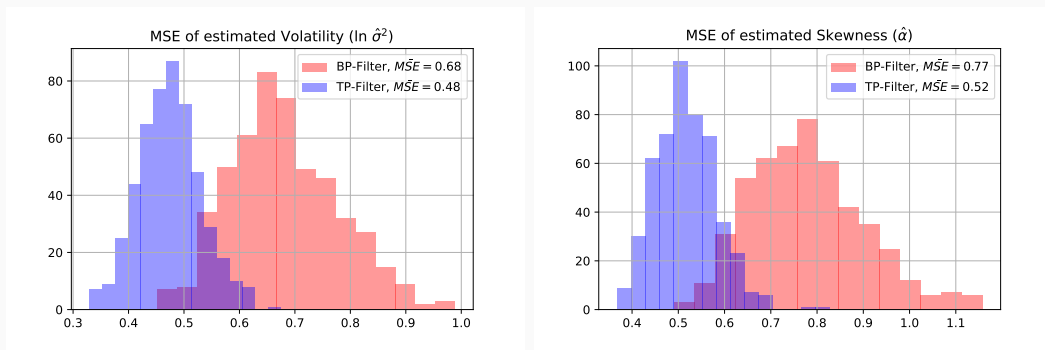


Figure 8: MSE of 500 filtered states obtained with the Bootstrap and the Tempered Particle Filter based on simulated data from Eq.1-3.

Posterior Distributions of Mean Equation

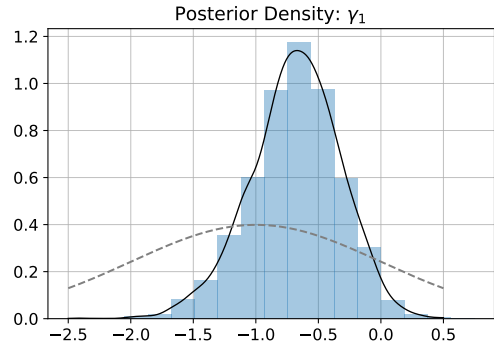
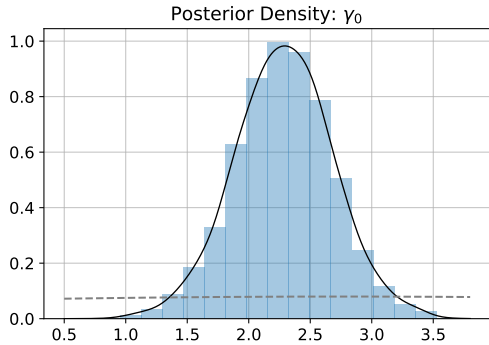


Figure 9: Posterior distributions of γ_1 and γ_2 obtained with the MCMC-Algorithm based on 20000 draws.

Tempered Particle Filter II

- The initial weights are given as

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

- The unnormalized weights at the n^{th} tempering step are given

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

with

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

- $\tilde{\Lambda}_{i,t}(\phi_n)$ introduces additional variance to \tilde{W}_i
- Since

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

additionally tempering the skewness results in less tempering iterations

Upside and Downside Entropy

$$\mathcal{L}_{\mathcal{M}_i,t}^U = - \int_{-\infty}^{\hat{F}_{\mathcal{M}_i,t}^{-1}(0.5)} (\log \hat{g}(y) - \log \hat{f}_{\mathcal{M}_i,t}(y)) \hat{f}_{\mathcal{M}_i,t}(y) dy \quad (8)$$

$$\mathcal{L}_{\mathcal{M}_i,t}^D = - \int_{\hat{F}_{\mathcal{M}_i,t}^{-1}(0.5)}^{\infty} (\log \hat{g}(y) - \log \hat{f}_{\mathcal{M}_i,t}(y)) \hat{f}_{\mathcal{M}_i,t}(y) dy \quad (9)$$

where $\hat{g}(y)$ is ML-estimator of the unconditional density and $\hat{f}_{\mathcal{M}_i,t}(y)$ is the estimated conditional density under model \mathcal{M}_i

- Relative measure of the divergence between two distributions in the upper and lower tails.
- If downside/upside entropy is high/low, more/less probability mass in the lower/upper tails of the conditional relative to the unconditional distribution

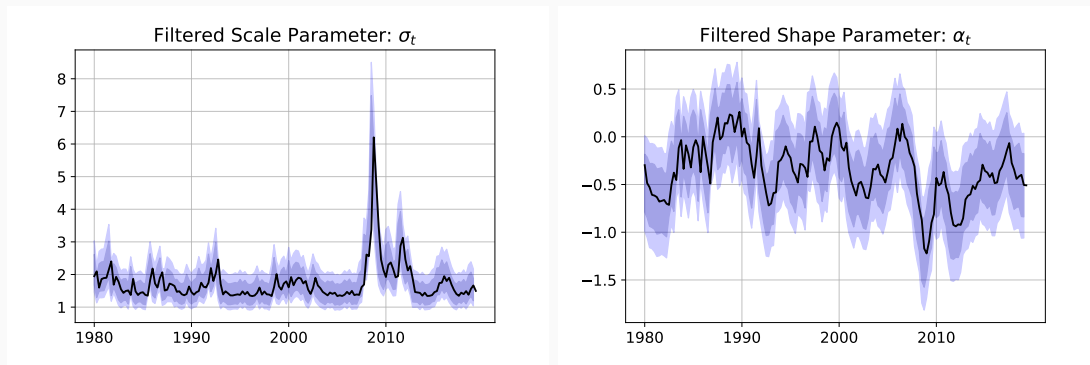


Figure 10: Filtered states given mean posterior values. The Tempered particle filter is tuned to target an Inefficiency Ratio with $\Delta_r = 0.01$, using 2 Mutation steps and $M = 10000$ particles. Shaded areas denote 68% and 90% credible sets.

- Significant variation of the second and third moment based on 68% and 90% credible sets.
- Inverse relationship less pronounced compared to US findings $\hat{\rho}[\hat{\sigma}_{t,eu}, \hat{\alpha}_{t,eu}] = -0.61532399$
- Skewness seems to be more centered around 0 with wider credible sets.

Growth at Risk: Conditional Densities in the Euro Area

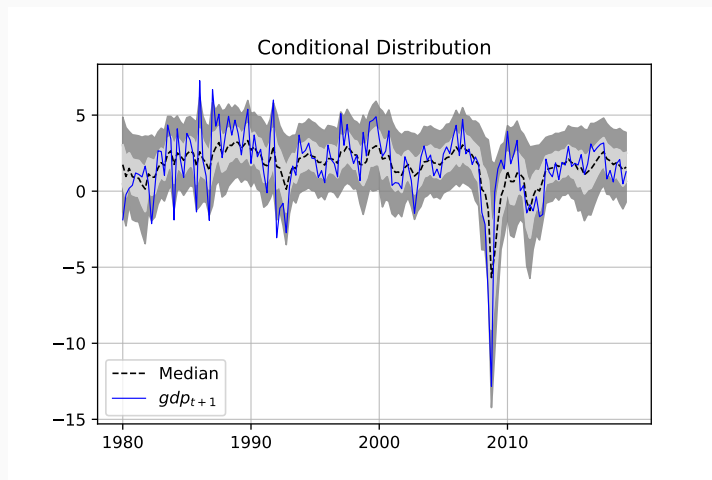


Figure 11: Estimated time-varying upper and lower 5% and 25% quantiles of the conditional forecasting densities.

Conditional forecasting densities of GDP growth conditional on Composite Indicator of Systemic Stress (CISS): Different behavior of the conditional forecasting densities is less pronounced in compared to the results of US data.