CIRJE-F-845

Bayesian Analysis of Time-Varying Quantiles Using a Smoothing Spline

Yuta Kurose
Graduate School of Economics, University of Tokyo
Yasuhiro Omori
University of Tokyo

March 2012

Bayesian analysis of time-varying quantiles using a smoothing spline

Yuta Kurose* Yasuhiro Omori†
*Graduate School of Economics, University of Tokyo
†Faculty of Economics, University of Tokyo

March 2012

Abstract

A smoothing spline is considered to propose a novel model for the time-varying quantile of the univariate time series using a state space approach. A correlation is further incorporated between the dependent variable and its one-step-ahead quantile. Using a Bayesian approach, an efficient Markov chain Monte Carlo algorithm is described where we use the multi-move sampler, which generates simultaneously latent time-varying quantiles. Numerical examples are provided to show its high sampling efficiency in comparison with the simple algorithm that generates one latent quantile at a time given other latent quantiles. Furthermore, using Japanese inflation rate data, an empirical analysis is provided with the model comparison.

Key words: Asymmetric double exponential distribution, Markov chain Monte Carlo, multi-move sampler, smoothing spline, state space approach, time-varying quantile.

1 Introduction

A time-varying quantile has been receiving attention recently, and various econometric models have been proposed. Tail quantiles are especially important for financial risk management or policy evaluation because they are useful to describe the extreme behavior of the dependent variable in serious events, such as the financial crisis. The Value at Risk (VaR) is an example of tail quantiles in financial time series, which is one of the well-known risk measures that are associated with an asset or a portfolio of many assets.

As discussed in the vast literature of financial econometrics, time-varying variances are found to exist in empirical studies of financial time series (see, e.g., Engle (1995), Shephard (2005)). As the variances of the dependent variables change over time, the corresponding quantiles vary over time. When we focus on the tail behavior of financial time series, it is necessary to describe the time-dependent structure that is appropriate for the tail quantile or τ-quantile (which is roughly defined as the value that will be exceeded by the dependent variable with probability τ) with τ close to zero or one. For the i.i.d. observations, the estimate is given by the 100τ-th percentile of the samples, and in the generalized linear model, it is given by the minimizer of some loss function (Koenker and Bassett (1978)). Based on Koenker and Bassett (1978), several models have been proposed for time-varying quantiles: Conditional Autoregressive Value at Risk (CAViaR) model (Engle and Manganelli (2004)), Quantile Autoregressive (QAR) model (Koenker and Xiao (2006)) and Dynamic Additive Quantile (DAQ) model (Gourieroux and Jasiak (2008)).

On the other hand, Koenker and Machado (1999) noted that solving the loss function of Koenker and Bassett (1978) is equivalent to obtaining the maximal likelihood estimate of the τ-quantile assuming some distribution for the dependent variable. For a Bayesian inference in the quantile regression, Yu and Moyeed (2001) took this approach to obtain the posterior distribution of the quantile and the parameters using the Markov chain Monte Carlo (MCMC) method. Gerlach, Chen, and Chan (2011) proposed a threshold-CAViaR model that extends the CAViaR model for the Bayesian analysis of time-varying quantiles.

In modeling time-varying quantiles, it is important to forecast the future tail behavior of the time series for risk management as well as to describe its past movement. Overfitting the model to the past dataset could result in a poor future forecasting for practical applications. Thus, recently, the backtesting procedure has been often implemented to investigate such a

forecasting performance, e.g., by checking that the proportion of observations that exceed the estimated one-step-ahead quantiles is equal to its expected value.

One approach to avoid overfitting to the past dataset is to make the time-varying quantile function so smooth that it produces stable predictions. In this paper, we use the smoothing spline for that purpose as discussed in De Rossi and Harvey (2009) and propose an efficient Bayesian estimation using the MCMC method in which we exploit a state space representation to apply a simulation smoother (de Jong and Shephard (1995), Durbin and Koopman (2002)) to generate the latent time-varying quantiles from the posterior distributions. The model is further extended to incorporate a correlation between the dependent variable and its one-step-ahead quantile.

The rest of this article is organized as follows. In Section 2, we propose the time-varying quantile model using the smoothing spline. Section 3 describes an efficient Bayesian estimation method for the proposed model using a multi-move sampling method. A single-move sampling method that is simple but inefficient is also described as a benchmark. We show that a normal variance-mean mixture representation of the measurement error leads us to exploit the efficient sampling method for the linear Gaussian state space model. Section 4 illustrates our estimation method using simulated data and shows that our MCMC algorithm is efficient. Section 5 applies the proposed time-varying quantile model to the inflation rate based on the domestic Corporate Goods Price Index (CGPI) of Japan. The backtesting procedure and the model comparison using DIC (Deviance Information Criterion) are conducted using our models and the CAViaR model. Section 6 concludes this paper.

2 Time-varying quantile model

2.1 Quantile regression model

Let y_t denote the dependent variable at time t $(t = 1, ..., n)$ whose distribution function is given by $F(y) = \Pr(y_t \leq y)$. For any fixed $0 < \tau < 1$, we define a τ-quantile as

$$\xi(\tau) := F^{-1}(\tau) = \inf\{y | F(y) \geq \tau\}, \tag{1}$$

and we define the loss function, called a "check function" as

$$\rho_\tau(u) = (\tau - \mathrm{I}(u < 0))(u), \tag{2}$$

where I$(\cdot)$ is an indicator function. Then, the expected loss,

$$\mathrm{E}(\rho_\tau(y_t - \xi(\tau))), \tag{3}$$

is minimized when $\xi(\tau)$ satisfies $F(\xi(\tau)) = \tau$. Using this loss function, Koenker and Bassett (1978) considered a quantile regression for i.i.d. observations assuming that $\xi(\tau) = \boldsymbol{x}'\boldsymbol{b}$, where $\boldsymbol{x}$ is a vector of explanatory variables and $\boldsymbol{b}$ is a corresponding regression coefficient vector (see, e.g., Koenker (2005) for the asymptotic property of the minimum loss estimator and the numerical method using a linear programing).

Yu and Moyeed (2001) assumed an asymmetric double exponential (or asymmetric Laplace) density that corresponds to the loss function and described an MCMC algorithm for the quantile regression using a Bayesian approach, where y_t is i.i.d. with the probability density function given $\xi(\tau)$,

$$\begin{aligned} f(y_t|\xi(\tau)) &= \frac{\tau(1-\tau)}{\lambda}\exp\Big(-\frac{1}{\lambda}\rho_\tau(y_t - \xi(\tau))\Big), \\ \rho_\tau(\epsilon_t) &= \begin{cases} (1-\tau)(-\epsilon_t) & (\epsilon_t < 0), \\ \tau\epsilon_t & (\epsilon_t \geq 0). \end{cases} \end{aligned} \tag{4}$$

and the first and second moments of $\epsilon_t = y_t - \xi(\tau)$ are (see, e.g., Kotz, Kozubowski, and Podgórski (2001))

$$\lambda\frac{-\tau^2 + (1-\tau)^2}{(1-\tau)\tau}, \quad 2\lambda^2\frac{\tau^3 + (1-\tau)^3}{(1-\tau)^2\tau^2}. \tag{5}$$

Note that $\xi(\tau)$, which maximizes the logarithm of this density function also minimizes the expected loss function, (3).

This paper extends the static model (4) to describe the time-varying quantiles, and we let ξ_t denote the time-varying τ-th quantile, where we suppress τ in a parenthesis and add a subscript t to emphasize that it depends on time t.

2.2 Time-varying quantile model using a smoothing spline

We assume that $\xi_t = h(t)$ changes slowly over time t and, hence, that $h(t)$ is a smooth function of t. That is, $h(t)$ is of a C^{m-1}-class, and its m-th derivative is square integrable and is the smoothing spline function, which minimizes

$$\sum_{t=1}^{n}\rho_\tau(y_t - h(t)) + \lambda_m\int[h^{(m)}(t')]^2dt' \tag{6}$$

for given m and λ_m. Furthermore, we assume that (i) h and its first $(m-1)$ derivatives at time $t=1$ follow the m-variate normal distribution with mean $\mathbf{0}_m$ and covariance matrix κE_m, where $\mathbf{0}_m$ is an m-dimensional zero vector, E_m is an identity matrix of size m, and κ is some known constant,

$$(h(1), h'(1), ..., h^{(m-1)}(1))' \sim \mathrm{N}(\mathbf{0}_m, \kappa E_m), \tag{7}$$

and that (ii)

$$h(t) = \sum_{j=1}^{m} \frac{(t-1)^{j-1}}{(j-1)!} h^{(j-1)}(1) + \sigma_\eta \int_1^t \frac{(t-s)^{m-1}}{(m-1)!} dW_s, \tag{8}$$

where W is a Wiener process.

Under these additional assumptions (i)(ii), if $\lambda_m = \lambda/(2\sigma_\eta^2)$, then the mode of the distribution of $(h(1), h(2), ..., h(n)|y_1, ..., y_n)$ converges to the solution of the smoothing spline problem as $\kappa \to \infty$ (De Rossi and Harvey (2009)). Noting that equation (8) can be represented in the following state space form (see, e.g., Wecker and Ansley (1983), Kohn and Ansley (1987)),

$$\boldsymbol{h}(t+1) = T\boldsymbol{h}(t) + \boldsymbol{\eta}(t), \boldsymbol{\eta}(t) \sim \mathrm{N}(\mathbf{0}_m, \sigma_\eta^2 Q), \tag{9}$$

$$\boldsymbol{h}(t) = (h(t), dh(t)/dt, ..., d^{m-1}h(t)/dt^{m-1})', \tag{10}$$

$$(T)_{ij} = \begin{cases} 1/(j-i)! & j \geq i, \\ 0 & j < i, \end{cases} \tag{11}$$

$$(Q)_{ij} = \frac{1}{(m-i)!(m-j)!(2m-i-j+1)}, \tag{12}$$

we propose a time-varying quantile model using the smoothing spline (TQSS model) in the state space representation with an asymmetric double exponential measurement error:

$$y_t = Z\boldsymbol{\xi}_t + \epsilon_t, \epsilon_t \sim \mathrm{aDE}_\tau(\lambda), \tag{13}$$

$$\boldsymbol{\xi}_{t+1} = T\boldsymbol{\xi}_t + \boldsymbol{\eta}_t, \boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}_m, \sigma_\eta^2 Q), \tag{14}$$

where

$$\boldsymbol{\xi}_t = (\xi_t, \tilde{\boldsymbol{\xi}}_t')', \quad \tilde{\boldsymbol{\xi}}_t = (d\xi_t/dt, ..., d^{m-1}\xi_t/dt^{m-1})', \quad (\xi_1, \xi_1^{(1)}, ..., \xi_1^{(m-1)})' \sim \mathrm{N}(\mathbf{0}_m, \kappa E_m),$$

$$Z = (1, \mathbf{0}_{m-1}').$$

Yue and Rue (2011) consider an additive mixed quantile regression model for longitudinal data with quantile functions including such a smooth function.

3 Bayesian estimation

3.1 Prior and posterior densities

For prior distributions of σ_η^2 and λ, we assume

$$\sigma_\eta^2 \sim \mathrm{IG}(\alpha_0/2, \beta_0/2), \quad \lambda \sim \mathrm{IG}(\alpha_0^*, \beta_0^*), \tag{15}$$

where $\mathrm{IG}(a, b)$ denotes an inverted gamma distribution with shape parameter a and scale parameter b. Let $I_t = \mathrm{I}(y_t - \xi_t < 0),\ t = 1, ..., n$. Then, the joint posterior density function is

$$\begin{aligned} f(\sigma_\eta^2, \lambda, \{\boldsymbol{\xi}_t\}_{t=1}^n | \{y_t\}_{t=1}^n) &\propto \prod_{t=1}^n f(y_t|\boldsymbol{\xi}_t, \lambda) \times f(\boldsymbol{\xi}_1) \prod_{t=1}^{n-1} f(\boldsymbol{\xi}_{t+1}|\boldsymbol{\xi}_t, \sigma_\eta^2) \times f(\sigma_\eta^2) f(\lambda) \\ &\propto \lambda^{-n} \exp\left(- \frac{\sum_{t=1}^n (\tau - I_t)(y_t - Z\boldsymbol{\xi}_t)}{\lambda} \right) \times \exp\left(- \frac{1}{2\kappa} \boldsymbol{\xi}_1' \boldsymbol{\xi}_1 \right) \\ &\times (\sigma_\eta^2)^{-m\frac{n-1}{2}} \exp\left(- \frac{1}{2} \sum_{t=1}^{n-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t)' (\sigma_\eta^2 Q)^{-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t) \right) \\ &\times (\sigma_\eta^2)^{-\left(\frac{\alpha_0}{2}+1\right)} \exp\left(- \frac{\beta_0}{2\sigma_\eta^2} \right) \times \lambda^{-(\alpha_0^*+1)} \exp\left(- \frac{\beta_0^*}{\lambda} \right). \end{aligned} \tag{16}$$

We implement the MCMC algorithm in five blocks:

1. Initialize $\sigma_\eta^2, \lambda, \{\boldsymbol{\xi}_t\}_{t=1}^n$.

2. Generate $\sigma_\eta^2 | \{y_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n \sim \mathrm{IG}(\alpha_1/2, \beta_1/2)$, where

$$\alpha_1 = \alpha_0 + m(n-1),\ \beta_1 = \beta_0 + \sum_{t=1}^{n-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t)' Q^{-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t). \tag{17}$$

3. Generate $\lambda | \{y_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n \sim \mathrm{IG}(\alpha_1^*, \beta_1^*)$, where

$$\alpha_1^* = \alpha_0^* + n,\ \beta_1^* = \beta_0^* + \sum_{t=1}^{n} (\tau - I_t)(y_t - Z\boldsymbol{\xi}_t). \tag{18}$$

4. For $t = 1, ..., n$, generate $\{\boldsymbol{\xi}_t\}_{t=1}^n | \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda$ as in Section 3.2.

5. Go to 2.

3.2 Generation of latent time-varying quantiles

3.2.1 Single-move sampling method

A simple sampling method for $\{\boldsymbol{\xi}_t\}_{t=1}^n$ is a single-move sampler that draws a single latent variable $\boldsymbol{\xi}_t$ at a time given the other $\boldsymbol{\xi}_t$'s and the parameters. The other method is a multi-move sampler that draws all of $\boldsymbol{\xi}_t$'s simultaneously. A single-move sampler is simpler than a multi-move sampler, but a multi-move sampler is known to be more efficient (de Jong and Shephard (1995)). As a benchmark, we describe the single-move sampling method as follows (see Appendix A.1 for details):

Step4. For $t = 1, ..., n$,

4.a Generate $I_t|\tilde{\boldsymbol{\xi}}_t, \{\boldsymbol{\xi}_{-t}\}, y_t, \sigma_\eta^2, \lambda$.

4.b Generate $\xi_t|I_t, \tilde{\boldsymbol{\xi}}_t, \{\boldsymbol{\xi}_{-t}\}, y_t, \sigma_\eta^2, \lambda$.

4.c Generate $\tilde{\boldsymbol{\xi}}_t|\xi_t, \{\boldsymbol{\xi}_{-t}\}, y_t, \sigma_\eta^2, \lambda$.

3.2.2 Efficient multi-move sampling

A simulation smoother, an efficient sampler for the state variables was proposed by de Jong and Shephard (1995) and by Durbin and Koopman (2002) for the linear Gaussian state space model. However, in the time-varying quantile model, the measurement error is non-Gaussian, and such a simulation smoother cannot be applied directly. For non-Gaussian measurement models, it is usually necessary to approximate the non-Gaussian likelihood by the Gaussian likelihood in the previous literature for the MCMC implementation (e.g., Shephard and Pitt (1997), Watanabe and Omori (2004), Kim, Shephard, and Chib (1998), Omori, Chib, Shephard, and Nakajima (2007)) and in our proposed model, the error distribution is asymmetric double exponential. Noting that it is a normal variance-mean mixture with a generalized inverted Gaussian distribution (e.g., Kotz, Kozubowski, and Podgórski (2001), Tsionas (2003), Kozumi and Kobayashi (2011) and Yue and Rue (2011)), we rewrite

$$\epsilon_t = av_t + b\sqrt{\lambda v_t}u_t, \tag{19}$$

$$a = \frac{1-2\tau}{\tau(1-\tau)},\ b^2 = \frac{2}{\tau(1-\tau)}, \tag{20}$$

where u_t is a standard normal variable and where v_t is an exponential variable with scale parameter (or mean) λ. Since the error distribution is normal conditionally on v_t, we can apply a simulation smoother for the linear Gaussian state space model.

Thus, to sample $\{\boldsymbol{\xi}_t\}_{t=1}^n$ efficiently, we rewrite the state space model (13)-(14) as follows:

$$y_t = Z\boldsymbol{\xi}_t + av_t + b\sqrt{\lambda v_t}u_t, u_t \sim \mathrm{N}(0,1), \;\; v_t \sim \mathrm{Exp}(\lambda), \tag{21}$$

$$\boldsymbol{\xi}_{t+1} = T\boldsymbol{\xi}_t + \boldsymbol{\eta}_t, \boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}_m, \sigma_\eta^2 Q). \tag{22}$$

The conditional joint posterior density of $\{\boldsymbol{\xi}_t\}_{t=1}^n$ and $\{v_t\}_{t=1}^n$ given $\{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda$ is

$$\begin{aligned}
&f(\{\boldsymbol{\xi}_t\}_{t=1}^n, \{v_t\}_{t=1}^n | \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda) \\
&\propto \prod_{t=1}^n f(y_t|\boldsymbol{\xi}_t, v_t, \lambda) \times \prod_{t=1}^n f(v_t|\lambda) \times f(\boldsymbol{\xi}_1) \prod_{t=1}^{n-1} f(\boldsymbol{\xi}_{t+1}|\boldsymbol{\xi}_t, \sigma_\eta^2) \\
&\propto \prod_{t=1}^n v_t^{-\frac{1}{2}} \exp\left\{ -\sum_{t=1}^n \frac{(y_t - Z\boldsymbol{\xi}_t - av_t)^2}{2b^2\lambda v_t} \right\} \times \exp\left\{ -\frac{\sum_{t=1}^n v_t}{\lambda} \right\} \\
&\times \exp\left(-\frac{1}{2\kappa}\boldsymbol{\xi}_1'\boldsymbol{\xi}_1 - \frac{1}{2}\sum_{t=1}^{n-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t)'(\sigma_\eta^2 Q)^{-1}(\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t) \right).
\end{aligned} \tag{23}$$

Thus, we generate v_t and $\boldsymbol{\xi}_t$ in two blocks:

Step4.

4.a' Generate $v_t|y_t, \boldsymbol{\xi}_t, \lambda \sim \mathrm{GIG}(1/2, \delta_t, \gamma)$[1] for $t = 1, ..., n$, where

$$\delta_t^2 = \frac{(y_t - Z\boldsymbol{\xi}_t)^2}{b^2\lambda}, \; \gamma^2 = \frac{2}{\lambda} + \frac{a^2}{b^2\lambda}. \tag{24}$$

4.b' Generate $\{\boldsymbol{\xi}_t\}_{t=1}^n | \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda, \{v_t\}_{t=1}^n$ using a simulation smoother (de Jong and Shephard (1995), Durbin and Koopman (2002))[2] .

Note that in Step 3 we generate $\lambda|\{y_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n$ and in Step 4.a' we generate $v_t|y_t, \boldsymbol{\xi}_t, \lambda$ for $t = 1, ..., n$ using collapsed Gibbs sampler (see, e.g., Chen, Shao, and Ibrahim (2000)).

[1] An efficient algorithm for random sampling from a generalized inverted Gaussian distribution ($\mathrm{GIG}(\nu, \delta, \gamma)$) is available from Dagpunar (1989), but the method does not work when δ or γ is too small. In that case, we adopt a rejection method using $\mathrm{Gamma}(\nu, \gamma^2/2)$ as a sampling method.

[2] For large m, the determinant of $\sigma_\eta^2 Q$ becomes so small that the round-off error for calculation is not negligible and that the covariance matrix of state ($\boldsymbol{\xi}_t$) estimation error can be non-positive definite. In such a case, we should use a square root filter by Morf and Kailath (1975) instead of an ordinary Kalman filter; see also Durbin and Koopman (2001).

3.3 Extension to the correlated errors

This subsection extends our model to describe a correlation between the dependent variable at time t and the latent time-varying quantile at time $(t+1)$. Based on the literature that addresses time-varying variances, the asymmetry or the leverage effect, which implies a decrease in the dependent variable at time t followed by an increase in the latent time-varying variance, is known to occur frequently in the empirical studies. Similarly, as discussed in Engle and Manganelli (2004), the time-varying quantile at time $(t+1)$ is also influenced by the observation at time t in their analysis of stock returns data. Thus, we incorporate a correlation between y_t and ξ_{t+1} using the state space representation as follows[3]:

$$y_t = Z\boldsymbol{\xi}_t + av_t + b\sqrt{\lambda v_t}u_t, u_t \sim \mathrm{N}(0,1), \;\; v_t \sim \mathrm{Exp}(\lambda), \tag{25}$$

$$\boldsymbol{\xi}_{t+1} = T\boldsymbol{\xi}_t + \boldsymbol{\eta}_t, \boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}_m, \sigma_\eta^2 Q), \tag{26}$$

where

$$\begin{pmatrix} b\sqrt{\lambda v_t}u_t \\ \boldsymbol{\eta}_t \end{pmatrix} \sim \mathrm{N}(\mathbf{0}, \Sigma_t), \; \Sigma_t = \left(\begin{array}{c|c} b^2\lambda v_t & \rho b\sigma_\eta\sqrt{\lambda v_t q_{11}} \quad \mathbf{0}'_{m-1} \\ \hline \rho b\sigma_\eta\sqrt{\lambda v_t q_{11}} \\ \mathbf{0}_{m-1} & \sigma_\eta^2 Q \end{array}\right), \tag{27}$$

$$|\Sigma_t| = b^2\lambda v_t \sigma_\eta^{2m}(|Q| - \rho^2 q_{11}|Q_{22}|), \tag{28}$$

q_{11} denotes the $(1,1)$ element of Q and Q_{22} denotes the matrix obtained by excluding the first row and column of Q. Assuming a uniform prior distribution for the correlation parameter, ρ[4],

$$\rho \sim \mathrm{U}(-c_m, c_m), \;\; c_m = \left\{|Q|/(q_{11}|Q_{22}|)\right\}^{\frac{1}{2}}, \tag{29}$$

(e.g., $c_2 = \frac{1}{2}, c_3 = \frac{1}{6}$) and assuming the same prior distributions for σ_η^2 and λ as in the previous subsection, the joint posterior density is

[3]Unlike the case without correlation, the mode of the distribution of $\{\xi_t\}_{t=1}^n|\{y_t\}_{t=1}^n$ may not necessarily converge to the solution to the smoothing spline problem (6).

[4]Given $\boldsymbol{\xi}_t$, the correlation coefficient between y_t and ξ_{t+1} is $\sqrt{\frac{(1-\tau)^2+\tau^2}{2\tau(1-\tau)}}\frac{\sqrt{\pi}}{2}\rho$. When $m=2$, the correlation coefficient of (y_t, ξ_{t+1}) is included in $(-0.208, 0.208)$ for $\tau = 0.1, 0.9$ and in $(-0.444, 0.444)$ for $\tau = 0.5$.

$$
\begin{aligned}
&f(\rho, \sigma_\eta^2, \lambda, \{v_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n | \{y_t\}_{t=1}^n) \\
&\propto f(\boldsymbol{\xi}_1) \prod_{t=1}^n f(y_t, \boldsymbol{\xi}_{t+1} | \boldsymbol{\xi}_t, \rho, \sigma_\eta^2, \lambda, v_t) \times \prod_{t=1}^n f(v_t|\lambda)\pi(\rho)\pi(\sigma_\eta^2)\pi(\lambda) \\
&\propto \exp\left\{ -\frac{\boldsymbol{\xi}_1' \boldsymbol{\xi}_1}{2\kappa} \right\} \times \prod_{t=1}^n |\Sigma_t|^{-\frac{1}{2}} \exp\left\{ -\sum_{t=1}^n \frac{1}{2} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - a v_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix}' \Sigma_t^{-1} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - a v_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix} \right\} \\
&\times \lambda^{-n} \exp\left\{ -\frac{\sum_{t=1}^n v_t}{\lambda} \right\} \times (\sigma_\eta^2)^{-(\frac{\alpha_0}{2}+1)} \exp\left\{ -\frac{\beta_0}{2\sigma_\eta^2} \right\} \times (\lambda)^{-(\alpha_0^*+1)} \exp\left\{ -\frac{\beta_0^*}{\lambda} \right\}. \qquad (30)
\end{aligned}
$$

We implement the MCMC algorithm in seven blocks:

1. Initialize $\rho, \sigma_\eta^2, \lambda, \{v_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n$.
2. Generate $\rho | \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda, \{v_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n$.
3. Generate $\sigma_\eta^2 | \{y_t\}_{t=1}^n, \rho, \lambda, \{v_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n$.
4. Generate $\lambda | \{y_t\}_{t=1}^n, \rho, \sigma_\eta^2, \{v_t\}_{t=1}^n, \{\boldsymbol{\xi}_t\}_{t=1}^n$.
5. Generate $v_t | y_t, \boldsymbol{\xi}_t, \rho, \sigma_\eta^2, \lambda$ for $t = 1, ..., n$.
6. Generate $\{\boldsymbol{\xi}_t\}_{t=1}^n | \{y_t\}_{t=1}^n, \rho, \sigma_\eta^2, \lambda, \{v_t\}_{t=1}^n$ using a simulation smoother as in the previous subsection.
7. Go to 2.

Generation of ρ. Let

$$
l(\rho) := -\frac{n}{2} \log(|Q| - \rho^2 q_{11} |Q_{22}|) - \sum_{t=1}^n \frac{1}{2} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - a v_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix}' \Sigma_t^{-1} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - a v_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix}, \qquad (31)
$$

which is the logarithm of the posterior density of ρ excluding the constant. To approximate the conditional posterior density by the truncated normal distribution, we use a Taylor expansion of the log posterior density of ρ to the second order around its mode $\hat{\rho}$ and let

$$
m_\rho = \hat{\rho} + s_\rho^2 l'(\hat{\rho}), \; s_\rho^2 = -1/l''(\hat{\rho}), \qquad (32)
$$

where $l'(\rho) = dl(\rho)/d\rho$ and $l''(\rho) = d^2 l(\rho)/d\rho^2$. Then, we propose a candidate $\rho^\dagger$ from the truncated normal distribution over the interval $(-c_m, c_m)$, $\mathrm{TN}_{(-c_m, c_m)}(m_\rho, s_\rho^2)$ and accept it with probability

$$
\min\left[1, \frac{\exp\{l(\rho^\dagger) - (\rho - m_\rho)^2/(2s_\rho^2)\}}{\exp\{l(\rho) - (\rho^\dagger - m_\rho)^2/(2s_\rho^2)\}}\right].
$$

Generation of σ_η^2. Propose a candidate $\sigma_\eta^{2\dagger} \sim \text{IG}(\alpha_1/2, \beta_1/2)$, where

$$\alpha_1 = \alpha_0 + mn,\ \beta_1 = \beta_0 + \sum_{t=1}^{n} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t)' Q^{-1} (\boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t), \tag{33}$$

and accept it with probability $\min\left[1, \exp\left(g(\sigma_\eta^{2\dagger}) - g(\sigma_\eta^2)\right)\right]$, where

$$g(\sigma_\eta^2) = -\frac{1}{2} \sum_{t=1}^{n} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix}' \Sigma_t^{-1} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix} + \frac{\beta_1}{2\sigma_\eta^2}. \tag{34}$$

Generation of λ. Propose a candidate $\lambda^\dagger \sim \text{IG}(\alpha_1^*, \beta_1^*)$, where

$$\alpha_1^* = \alpha_0^* + \frac{3}{2}n,\ \beta_1^* = \beta_0^* + \sum_{t=1}^{n} \frac{(y_t - Z\boldsymbol{\xi}_t - av_t)^2}{2b^2 v_t} + \sum_{t=1}^{n} v_t, \tag{35}$$

and accept it with probability $\min\left[1, \exp\left(g(\lambda^\dagger) - g(\lambda)\right)\right]$, where

$$g(\lambda) = -\frac{1}{2} \sum_{t=1}^{n} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix}' \Sigma_t^{-1} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix} + \frac{\beta_1^*}{\lambda}. \tag{36}$$

Generation of v_t. Propose a candidate $v_t^\dagger \sim \text{GIG}(1/2, \delta_t, \gamma)$, where

$$\delta_t^2 = \frac{(y_t - Z\boldsymbol{\xi}_t)^2}{b^2\lambda},\ \gamma^2 = \frac{2}{\lambda} + \frac{a^2}{b^2\lambda}, \tag{37}$$

and accept it with probability $\min\left[1, \exp\left(g(v_t^\dagger) - g(v_t)\right)\right]$, where

$$g(v_t) = -\frac{1}{2} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix}' \Sigma_t^{-1} \begin{pmatrix} y_t - Z\boldsymbol{\xi}_t - av_t \\ \boldsymbol{\xi}_{t+1} - T\boldsymbol{\xi}_t \end{pmatrix} + \frac{1}{2}(v_t^{-1}\delta_t^2 + v_t\gamma^2). \tag{38}$$

4 Illustrative examples using simulated data

This section illustrates our proposed time-varying quantile models using simulated data. We consider the case $m = 2$ and assume that the variance parameter for the initial state κ is 100. The sensitivity analysis for the selection of κ is also investigated.

First, we show the high efficiency of our multi-move sampling method in comparison with the single-move sampling method using the TQSS model given by (13)-(14). Using the following parameters based on our empirical studies in Section 5,

$$\sigma_\eta^2 = 4.0 \times 10^{-3}, \quad \lambda = 3.5 \times 10^{-2}, \quad \text{for } \tau = 0.1,$$
$$\sigma_\eta^2 = 1.0 \times 10^{-4}, \quad \lambda = 4.0 \times 10^{-2}, \quad \text{for } \tau = 0.9,$$

we generate three hundred observations ($n = 300$) for each τ. For prior distributions, we assume

$$\sigma_\eta^2 \sim \mathrm{IG}(0.1, 0.00005), \quad \lambda \sim \mathrm{IG}(0.1, 0.1).$$

Using the single-move sampler (Section 3.2.1), we generate 600,000 (450,000) MCMC samples after discarding the first 1,000 (1,000) samples as the burn-in period for $\tau = 0.1$ ($\tau = 0.9$). Also, the multi-move sampler (Section 3.2.2) is used to generate 30,000 (15,000) MCMC samples after discarding the first 1,000 (1,000) samples as the burn-in period for $\tau = 0.1$ ($\tau = 0.9$).

Table 1: TQSS model ($\tau = 0.1$).
Posterior means, standard deviations, 95% credible intervals and inefficiency factors (IF).

		True	Mean	Stdev	95% interval	IF
single-move	$\sigma_\eta^2 \times 10^3$	4	3.398	0.743	[2.206, 5.091]	133
	$\lambda \times 10^2$	3.5	3.296	0.215	[2.900, 3.741]	46
multi-move	$\sigma_\eta^2 \times 10^3$	4	3.374	0.766	[2.175, 5.169]	31
	$\lambda \times 10^2$	3.5	3.302	0.214	[2.908, 3.748]	2

Table 2: TQSS model ($\tau = 0.9$).
Posterior means, standard deviations, 95% credible intervals and inefficiency factors (IF).

		True	Mean	Stdev	95% interval	IF
single-move	$\sigma_\eta^2 \times 10^4$	1	0.976	0.343	[0.499, 1.817]	530
	$\lambda \times 10^2$	4	4.018	0.241	[3.572, 4.517]	51
multi-move	$\sigma_\eta^2 \times 10^4$	1	0.939	0.317	[0.495, 1.700]	44
	$\lambda \times 10^2$	4	4.025	0.242	[3.578, 4.537]	2

Tables 1 and 2 report the true values, posterior means, posterior standard deviations, 95% credible intervals and estimates of inefficiency factors (IF). The inefficiency factor is defined as $1 + 2\sum_{g=1}^{\infty} \rho(g)$, where $\rho(g)$ is the sample autocorrelation at lag g. This is interpreted as the ratio of the numerical variance of the posterior mean from the chain to the variance of the posterior mean from hypothetical uncorrelated draws. The smaller the inefficiency factor becomes, the closer the MCMC sampling is to the uncorrelated sampling.

The posterior means are all close to the true values, which suggests that our proposed algorithms work well.

The inefficiency factors for the single-move sampler are much larger than those for the

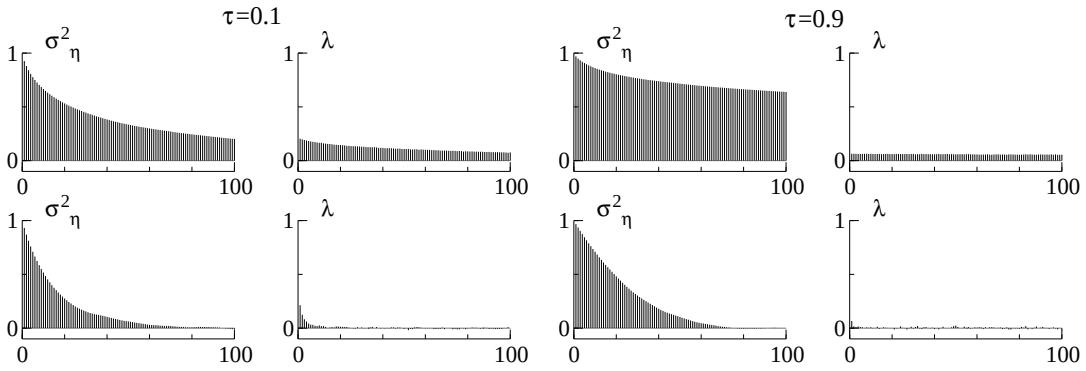

Figure 1: TQSS model. Sample autocorrelation functions of MCMC samples for the single-move sampler (top) and the multi-move sampler (bottom).

multi-move sampler, which suggests that our multi-move sampling method is highly efficient compared with the single-move sampling method. Figure 1 shows the sample autocorrelation functions where autocorrelations decay slowly for the single-move sampler and vanish quickly for the multi-move sampler, which also implies the efficiency of our multi-move sampler.

Next, we illustrate our estimation method for the TQSS model with correlations and investigate the sensitivity analysis with respect to the selection of the initial variance parameter κ.

Setting the parameters,

$$\rho = -0.35, \quad \sigma_\eta^2 = 4.0 \times 10^{-3}, \quad \lambda = 3.0 \times 10^{-2}, \quad \text{for } \tau = 0.1,$$
$$\rho = -0.33, \quad \sigma_\eta^2 = 1.5 \times 10^{-4}, \quad \lambda = 4.0 \times 10^{-2}, \quad \text{for } \tau = 0.9,$$

we generate three hundred observations ($n = 300$) for each $\tau = 0.1, 0.9$. For prior distributions, we assume a uniform distribution for ρ, $\rho \sim \mathrm{U}(-1/2, 1/2)$, and the same distributions for σ_η^2 and λ. We generate 1,200,000 (600,000) MCMC samples after discarding 1,000 (1,000) samples as the burn-in period for $\tau = 0.1$ ($\tau = 0.9$) using the multi-move sampler (Section 3.3) with $\kappa = 10, 100$ and 1000.

The estimation results are summarized in Table 3. For all κ, the posterior means of the parameters are close to the true values, which suggests that our sampler works well. The inefficiency factors are larger overall than those for the model without correlations due to the inefficiency of sampling the new parameter ρ. The sampling method for ρ needs to be improved but will be left for future work.

Taking into account the posterior standard deviations, the estimation results are robust with respect to the selection of κ. Figure 2 shows the estimated posterior densities of the parameters for $\kappa = 10, 100$ and 1000. They also show the robustness of the estimation results with respect to the selection of κ.

Table 3: TQSS model with correlations ($\tau = 0.1, 0.9$). Posterior means, standard deviations, 95% credible intervals and inefficiency factors for $\kappa = 10, 100$ and 1000.

		True	Mean	Stdev	95% interval	IF
$\tau = 0.1$	ρ	-0.35	-0.128	0.266	[-0.488, 0.431]	475
($\kappa = 100$)	$\sigma_\eta^2 \times 10^3$	4	3.362	0.761	[2.167, 5.125]	73
	$\lambda \times 10^2$	3	3.006	0.204	[2.629, 3.429]	42
	ρ	-0.35	-0.127	0.264	[-0.484, 0.434]	487
($\kappa = 10$)	$\sigma_\eta^2 \times 10^3$	4	3.364	0.756	[2.166, 5.106]	66
	$\lambda \times 10^2$	3	3.006	0.202	[2.633, 3.424]	40
	ρ	-0.35	-0.111	0.271	[-0.484, 0.443]	496
($\kappa = 1000$)	$\sigma_\eta^2 \times 10^3$	4	3.354	0.758	[2.160, 5.103]	63
	$\lambda \times 10^2$	3	3.011	0.203	[2.636, 3.430]	41
$\tau = 0.9$	ρ	-0.33	-0.047	0.283	[-0.484, 0.459]	461
($\kappa = 100$)	$\sigma_\eta^2 \times 10^4$	1.5	1.875	0.617	[0.995, 3.363]	109
	$\lambda \times 10^2$	4	3.768	0.228	[3.346, 4.241]	17
	ρ	-0.33	-0.038	0.279	[-0.473, 0.456]	464
($\kappa = 10$)	$\sigma_\eta^2 \times 10^4$	1.5	1.892	0.626	[1.009, 3.433]	105
	$\lambda \times 10^2$	4	3.768	0.228	[3.347, 4.241]	13
	ρ	-0.33	-0.052	0.290	[-0.485, 0.470]	509
($\kappa = 1000$)	$\sigma_\eta^2 \times 10^4$	1.5	1.877	0.621	[0.991, 3.378]	122
	$\lambda \times 10^2$	4	3.766	0.228	[3.345, 4.237]	14

5 Empirical study

5.1 Data

Economic agents are known to consider not only the mean of the response distribution but also many aspects of the distribution in their decision making processes. In macroeconomics, for example, the upper tail of the distribution of inflation rate is usually one of the greatest concerns to the central bank because monetary policy measures are deployed that focus on (violent) inflation in an attempt to keep the price movement stable. On the other hand, the lower tail of the distribution of inflation rate would attract increasing attention of those advanced countries who face the risk of deflation, which may cause a serious recession.

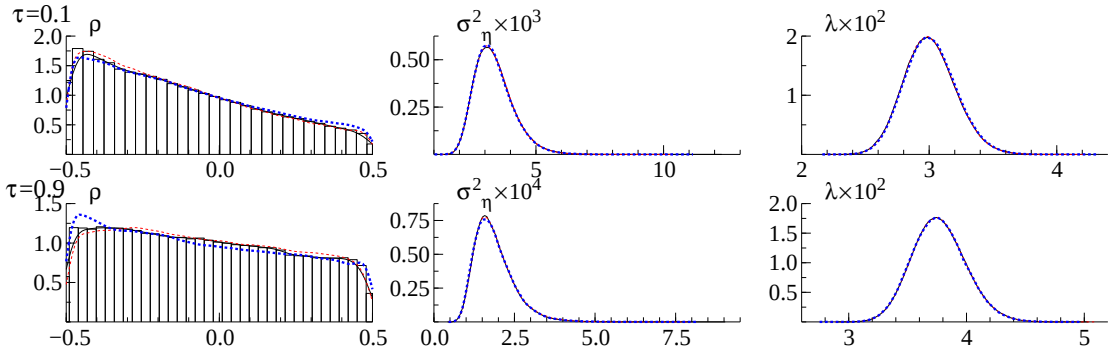

Figure 2: Time-varying quantile model with correlations ($\tau = 0.1, 0.9$). Estimated posterior densities for $\kappa = 10$ (thin dotted line), $\kappa = 100$ (solid line with histogram), $\kappa = 1000$ (thick dotted line).

Thus, this section applies our proposed model to the upper and the lower tails of the distribution of the inflation rate of Japan, using the rate of change for the domestic Corporate Goods Price Index (CGPI) excluding the consumption tax for all commodities of Japan (reported by Bank of Japan). The rate of change is calculated as $y_t = 100 \times (\log p_t - \log p_{t-1})$, where p_t is the CGPI at time t. We consider the following two sample periods:

$$\text{Period (I)}: \text{ February, 1985 – June, 2008 (281 months) and}$$

$$\text{Period (II)}: \text{ February, 1985 – January, 2010 (300 months).}$$

Period (I) is set before Lehman Brothers filed for Chapter 11 bankruptcy protection (September 15, 2008) to eliminate the event's significant negative impact on the economy. We consider two time-varying quantiles with $\tau = 0.1, 0.9$ as mentioned above. Table 4 shows the summary statistics for the inflation rate based on the CGPI of Japan (y_t). The sample mean of the inflation rate for Period (II) is slightly lower than that for Period (I), while the maximum and the absolute value of the minimum for Period (II) are much larger than those for Period (I). This suggests the existence of the greater uncertainty in the inflation rate during Period (II).

5.2 Estimation results

Using the same prior distributions for the parameters as in Section 4, $m = 2$ and $\kappa = 100$, we implement the MCMC algorithm to conduct a Bayesian inference on parameters of interest.

Table 4: Summary statistics for the inflation rate based on CGPI of Japan (y_t).

	Nob	Mean	Stdev	Max	Min
Period (I) [1985.2-2008.6]	281	-0.038	0.279	1.201	-0.975
Period (II) [1985.2-2010.1]	300	-0.059	0.361	2.160	-1.989

We generate 60,000 (30,000, 30,000) MCMC samples from the posterior distributions of the parameters in the model without correlations and generate 60,000 (240,000, 240,000) MCMC samples from the posterior distributions of the parameters in the model with correlations, after discarding the first 1,000 (1,000, 1,000) samples as the burn-in period for $\tau = 0.1$ ($\tau = 0.5$, $\tau = 0.9$).

Table 5: Period (I). TQSS model.

		Mean	Stdev	95% interval	IF
$\tau = 0.1$	$\sigma_\eta^2 \times 10^4$	3.469	3.355	[0.523, 12.66]	210
	$\lambda \times 10^2$	3.402	0.261	[2.902, 3.931]	77
$\tau = 0.5$	$\sigma_\eta^2 \times 10^5$	5.338	3.143	[1.703, 13.77]	119
	$\lambda \times 10^2$	7.626	0.487	[6.723, 8.638]	6
$\tau = 0.9$	$\sigma_\eta^2 \times 10^4$	1.424	0.695	[0.570, 3.214]	87
	$\lambda \times 10^2$	3.159	0.204	[2.783, 3.579]	5

Table 6: Period (I). TQSS model with correlations.

		Mean	Stdev	95% interval	IF
$\tau = 0.1$	ρ	-0.253	0.211	[-0.492, 0.293]	228
	$\sigma_\eta^2 \times 10^4$	4.299	3.268	[0.629, 13.19]	259
	$\lambda \times 10^2$	3.316	0.263	[2.825, 3.855]	115
$\tau = 0.5$	ρ	0.0192	0.294	[-0.473, 0.476]	391
	$\sigma_\eta^2 \times 10^5$	4.970	2.876	[1.444, 12.26]	223
	$\lambda \times 10^2$	7.640	0.485	[6.745, 8.642]	18
$\tau = 0.9$	ρ	-0.114	0.278	[-0.485, 0.455]	389
	$\sigma_\eta^2 \times 10^4$	1.473	0.759	[0.562, 3.435]	196
	$\lambda \times 10^2$	3.150	0.205	[2.770, 3.576]	26

The estimation results for Period (I) are given in Tables 5 and 6 for the two models ($\tau = 0.1, 0.5, 0.9$). The parameter estimates are quite similar for both models taking into account the posterior standard deviations. The estimate of the variance parameter of the

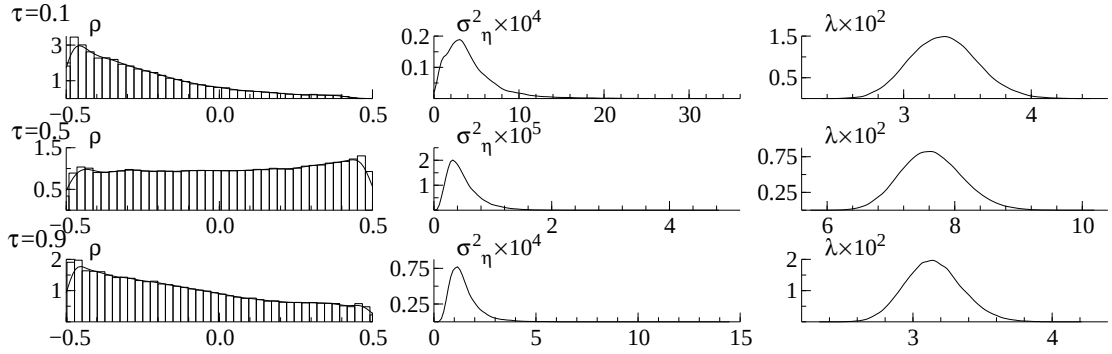

Figure 3: Period (I). TQSS model with correlations.

state equation σ_η^2 for $\tau = 0.1$ is much larger than that for $\tau = 0.9$, indicating that the lower tail quantile is more uncertain than the upper tail quantile. The posterior means of the correlations are negative for $\tau = 0.1, 0.9$, but they are not credible since their 95% credible intervals include zeros. The estimated posterior densities are shown in Figure 3 in which ρ seems to have a large posterior probability on negative values especially for $\tau = 0.1$[5].

Furthermore, Figure 4 shows the time series plot of the posterior means of ξ_t and y_t for $\tau = 0.1, 0.5, 0.9$. The estimated quantile posterior means vary smoothly and capture the changes in the level and the magnitude of the inflation rate over the sample period. The estimates for the two TQSS models with and without correlations are found to be quite similar (solid and thick dotted lines, respectively).

Table 7: Period (II). TQSS model.

		Mean	Stdev	95% interval	IF
$\tau = 0.1$	$\sigma_\eta^2 \times 10^3$	5.418	2.255	[1.570, 10.48]	114
	$\lambda \times 10^2$	3.274	0.294	[2.758, 3.913]	67
$\tau = 0.5$	$\sigma_\eta^2 \times 10^5$	4.226	3.480	[0.996, 14.03]	153
	$\lambda \times 10^2$	9.717	0.600	[8.607, 10.94]	12
$\tau = 0.9$	$\sigma_\eta^2 \times 10^4$	1.425	0.813	[0.525, 3.288]	103
	$\lambda \times 10^2$	4.084	0.254	[3.619, 4.612]	7

Tables 7 and 8 show the estimation results for Period (II), and the estimated posterior densities are shown in Figure 5. The results between the two models are quite similar, and

[5]Period (I). The posterior mean of $\text{corr}(y_t, \xi_{t+1})$ given $\boldsymbol{\xi}_t$ is -0.105 (0.0170, -0.0473) for $\tau = 0.1$ (0.5, 0.9).

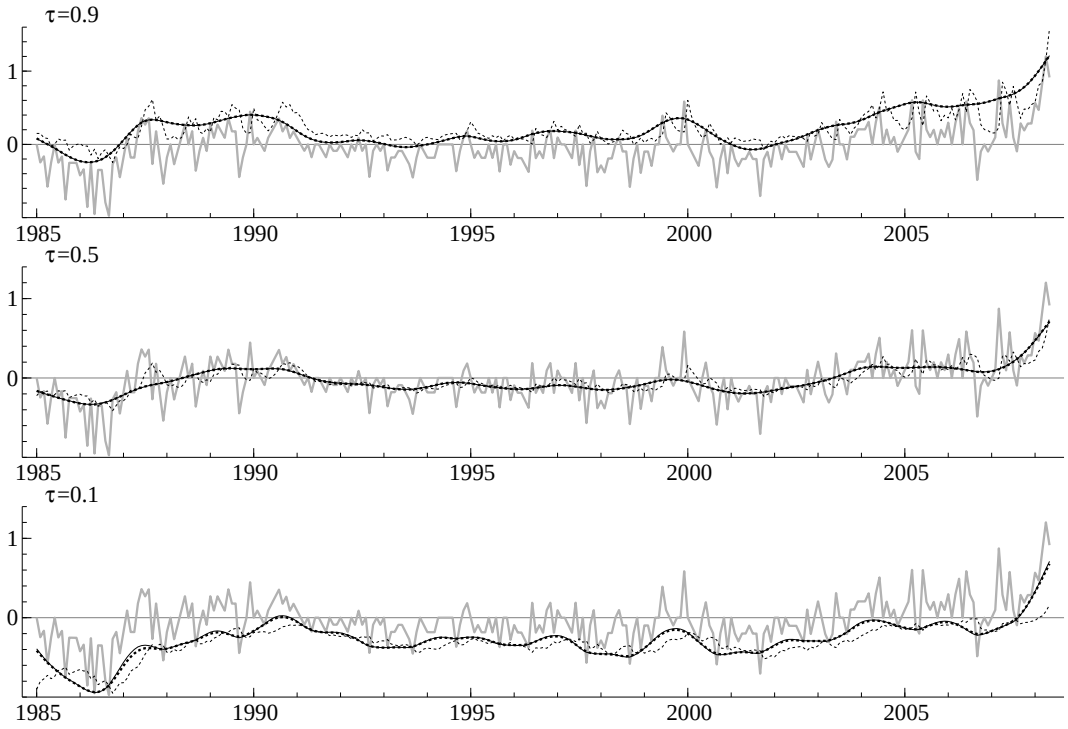

Figure 4: Period (I). Time series plot of the inflation rate (gray line), posterior means of time-varying quantiles using TQSS models without correlations (thick dotted line), with correlations (solid line) and CAViaR model (thin dotted line).

Table 8: Period (II). TQSS model with correlations.

		Mean	Stdev	95% interval	IF
$\tau = 0.1$	ρ	-0.351	0.143	[-0.496, 0.031]	219
	$\sigma_\eta^2 \times 10^3$	5.679	2.114	[2.157, 10.21]	192
	$\lambda \times 10^2$	3.159	0.289	[2.631, 3.771]	133
$\tau = 0.5$	ρ	0.0410	0.293	[-0.474, 0.485]	389
	$\sigma_\eta^2 \times 10^5$	4.397	4.312	[1.005, 16.21]	374
	$\lambda \times 10^2$	9.722	0.602	[8.605, 10.96]	39
$\tau = 0.9$	ρ	-0.092	0.276	[-0.486, 0.448]	391
	$\sigma_\eta^2 \times 10^4$	1.436	0.683	[0.528, 3.166]	215
	$\lambda \times 10^2$	4.077	0.255	[3.606, 4.604]	25

the posterior means of σ_η^2 for $\tau = 0.1$ are found to be much larger than those for $\tau = 0.9$ as for Period (I). We note that the estimate of σ_η^2 for $\tau = 0.1$ for Period (II) is ten times as large as the corresponding estimate for Period (I). This is because the inflation rate fluctuated

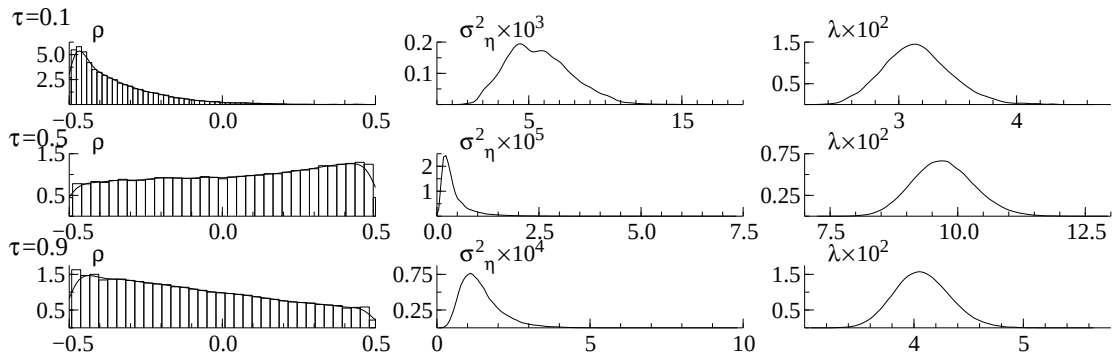

Figure 5: Period (II). TQSS model with correlations.

largely around September 2008 for Period (II). It suggests that the deflationary impact in Japan was more serious than the inflation during this period[6].

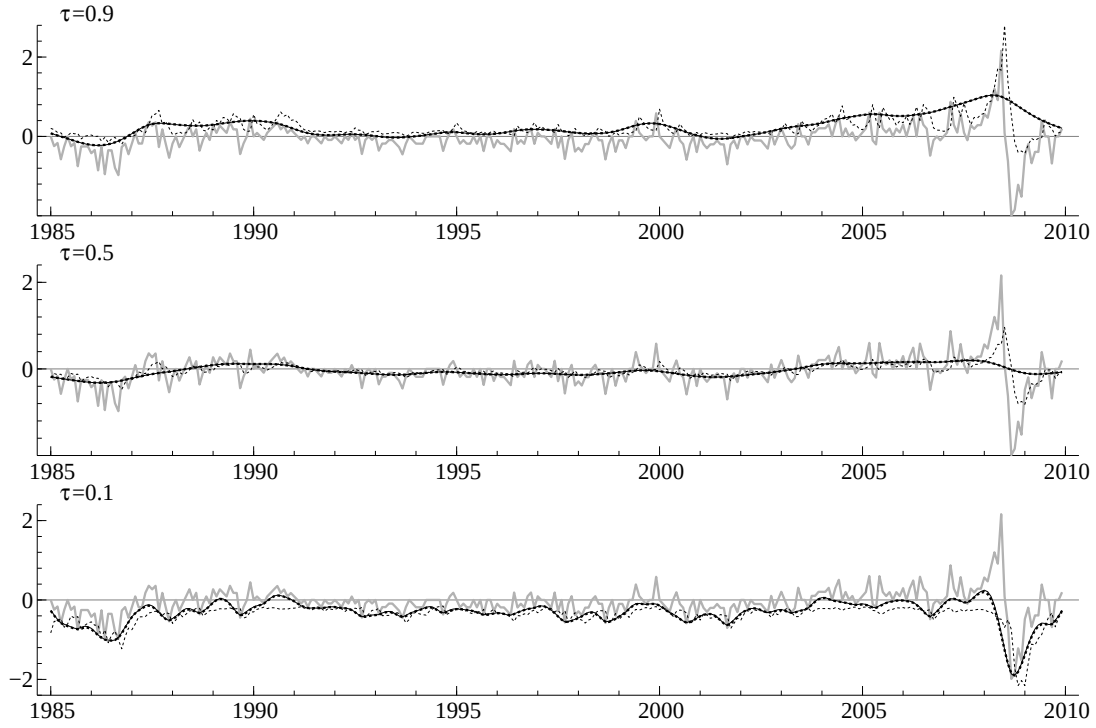

Figure 6: Period (II). Time series plot of the inflation rate (gray line), posterior means of time-varying quantiles using TQSS models without correlations (thick dotted line), with correlations (solid line) and CAViaR model (thin dotted line).

[6]Period (II). The posterior mean of $\mathrm{corr}(y_t, \xi_{t+1})$ given $\boldsymbol{\xi}_t$ is -0.146 (0.0363, -0.0382) for $\tau = 0.1$ (0.5, 0.9).

Figure 6 shows the time series plot of the posterior means of ξ_t and y_t for $\tau = 0.1, 0.5, 0.9$. The estimated quantile posterior means are similar for the two models, analogous to Figure 4 for Period (I). However, in contrast to the plot for Period (I), there is a sharp downward spike around the end of Period (II) for $\tau = 0.1$, corresponding to the large estimate of the state variance parameter σ_η^2.

As a benchmark model for the time-varying quantiles, we also estimate the CAViaR (Conditional Autoregressive Value at Risk) model by Engle and Manganelli (2004). The asymmetric slope type model is chosen among models in the CAViaR class because it is able to capture the effect of asymmetry in time-varying quantiles (see Appendix A.2 for more details). The posterior means of ξ_t based on the CAViaR (asymmetric slope) model are also shown in Figures 4 and 6 (thin dotted lines). The plots seem to be rough and sensitive to the change in the inflation rate compared with those estimates from our proposed models based on the smoothing spline.

5.3 Model Comparison

This subsection conducts a model comparison of our proposed models and the CAViaR model based on the backtesting (Kupiec (1995)) and the DIC (Spiegelhalter, Best, Carlin, and van der Linde (2002)).

Backtesting. Whether the future observation exceeds the τ-quantile ($(1-\tau)$-VaR) is an important issue in terms of decision making for economic policy and for financial risk management. To evaluate such a forecasting performance, we conduct the likelihood test by Kupiec (1995), which is often used for backtesting for VaR in finance. First, we fix $n_0 (< n)$ and set $s = 0$. Then, we (i) estimate the parameters and the latent quantiles using observations $\{y_t\}_{t=s+1}^{s+n_0}$ and (ii) compute the posterior mean of the predictive distribution of the one-step-ahead quantile, ξ_{s+n_0+1}. We repeat (i) and (ii) for $s = 0, 1, ..., n - n_0$ and compute the number of times N when the posterior mean $\hat{\xi}_{s+n_0+1}$ exceeds the observed data y_{s+n_0+1} for $s = 0, 1, ..., n - n_0$. If the null hypothesis that $\Pr(y_t < \xi_t) = \tau$ is true (and the probability is independent for each t),

$$2\left\{ \log\left(\left(\frac{N}{n-n_0}\right)^N \left(1 - \frac{N}{n-n_0}\right)^{n-n_0-N} \right) - \log(\tau^N (1-\tau)^{n-n_0-N}) \right\} \tag{39}$$

is asymptotically distributed as $\chi^2(1)$[7] .

[7]Since the time-varying quantiles are not independent of one another, the asymptotic result may not hold.

Table 9: p-values of the likelihood test.

	Model	Period (I)	Period (II)
$\tau = 0.1$	TQSS	0.021	0.000
	TQSSC	0.021	0.000
	CAViaR	0.417	0.743
τ=0.9	TQSS	0.045	0.032
	TQSSC	0.045	0.032
	CAViaR	0.009	0.001

TQSSC: TQSS model with correlations.

Table 9 shows the p-values of the one-sided likelihood ratio tests using $n_0 = 200$ observations for both Periods (I) and (II) ($n = 281$ for Period (I) and $n = 300$ for Period (II)). The number of the MCMC iterations and the prior distributions for each estimation are the same as those of the previous subsection. For Period (I), the null hypothesis is rejected for the CAViaR model ($\tau = 0.9$) since its p-value is smaller than 0.01. For Period (II), we reject the null hypotheses for the TQSS models ($\tau = 0.1$) and for the CAViaR model ($\tau = 0.9$). Thus, for the upper tail quantile ($\tau = 0.9$), our proposed models show good forecasting performances with respect to VaR, while the CAViaR model fails for both periods. On the other hand, for the lower tail quantile ($\tau = 0.1$), the CAViaR model performs well for both periods, while the performance of our proposed models depends on the sample period.

Model selection based on DIC. The Deviance Information Criterion (DIC) is used as a Bayesian measure of fit or adequacy and is defined as

$$\text{DIC} = \text{E}_{\boldsymbol{\theta}|Y_n}[D(\boldsymbol{\theta})] + p_D, \tag{40}$$

where $D(\boldsymbol{\theta}) = -2\log f(Y_n|\boldsymbol{\theta})$, $p_D = \text{E}_{\boldsymbol{\theta}|Y_n}[D(\boldsymbol{\theta})] - D(\text{E}_{\boldsymbol{\theta}|Y_n}[\boldsymbol{\theta}])$ represents model complexity as a penalty, $Y_n = \{y_t\}_{t=1}^n$ and $\boldsymbol{\theta}$ denotes the parameters. We estimate $\text{E}_{\boldsymbol{\theta}|Y_n}[D(\boldsymbol{\theta})]$ using the sample analogue $\overline{D(\boldsymbol{\theta}_{(d)})} = \frac{1}{d^*}\sum_{d=1}^{d^*} D(\boldsymbol{\theta}_{(d)})$, where $\boldsymbol{\theta}_{(d)}$s are resampled from the posterior distribution. We set d^* equal to 600 ($\tau = 0.1$) and 300 ($\tau = 0.9$) for the TQSS model, 600 ($\tau = 0.1$) and 2,400 ($\tau = 0.9$) for the TQSS model with correlations and 240 for the CAViaR model. Because we need to compute $D(\boldsymbol{\theta})$ numerically, we use the particle filter (e.g., Doucet, de Freitas, and Gordon (2001)), where we set the number of particles $M = 10,000$

The implications of the likelihood test should be used with caution.

(see Appendix A.3 for details). The numerical standard error of the estimate is obtained by repeating the particle filter forty times.

Table 10: DIC (standard errors in parentheses).

	Model	Period (I)		Period (II)	
$\tau = 0.1$	TQSS	141.16	(0.19)	329.86	(2.34)
	TQSSC	138.31	(0.25)	249.42	(0.68)
	CAViaR	141.69	(0.14)	283.44	(0.10)
$\tau = 0.9$	TQSS	92.93	(0.24)	259.29	(1.52)
	TQSSC	91.51	(0.19)	240.72	(2.26)
	CAViaR	102.84	(0.09)	190.11	(0.10)

TQSSC: TQSS model with correlations

Table 10 shows the sample means of forty DICs for each model with the standard errors in parentheses. The DICs of the TQSS model with correlations are the smallest and outperform the other competing models for $\tau = 0.1, 0.9$ for Period (I) and for $\tau = 0.1$ for Period (II). However, for $\tau = 0.9$ for Period (II), the CAViaR model outperforms the two TQSS models. This is partly because the quantile estimates of the CAViaR model follow the fluctuations of the inflation rate more quickly than those of the TQSS models as seen in Figure 6.

In summary, for Period (I), as illustrated in Figure 4, the trajectories of the estimated quantiles of the two TQSS models are smooth to obtain good forecasting performances for the VaR. In addition, the TQSS model with correlation attains the smallest DIC. However, with respect to Period (II), there is no model for which the null hypothesis of the backtesting is accepted using both $\tau = 0.1$ and 0.9. The DIC suggests that the different model should be used depending on τ.

6 Conclusion

This article proposed the novel smoothing spline model for time-varying quantiles. Taking a Bayesian approach, the efficient MCMC algorithm is described using a normal variance-mean mixture representation of the measurement error term where we exploit a simulation smoother for the linear Gaussian state space model. Its high efficiency is illustrated using simulated data in comparison with the single-move sampler. The model is extended to incorporate a correlation between the dependent variable and its one-step-ahead quantile. Furthermore, in

comparison with the CAViaR model, our method is shown to perform well regarding both one-ahead predictions and goodness-of-fit in the analysis of Japanese inflation rate.

Acknowledgements

The authors would like to thank Professor Michael McAleer, Professor Ryozo Miura, Professor Kosuke Oya, Professor Akimichi Takemura, Professor Masanobu Taniguchi, Professor Toshiaki Watanabe and two anonymous referees for their valuable comments. This work is supported by the Research Fellowship (DC1) from the Japan Society for the Promotion of Science and by the Grants-in-Aid for Scientific Research (A) 21243018 from the Japanese Ministry of Education, Science, Sports, Culture and Technology.

The computational results are obtained using Ox version 5.10 (see Doornik (2007)).

Appendix

A.1 Generation of $\{\boldsymbol{\xi}_t\}$ using a single-move sampler

In Section 3.2.1, the joint posterior density of $\boldsymbol{\xi}_t, I_t|\{y_t\}_{t=1}^n, \{\boldsymbol{\xi}_{-t}\}, \sigma_\eta^2, \lambda$ is given by

$$f(\boldsymbol{\xi}_t, I_t|\{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda) \propto \exp\left(-\frac{1}{2}\{\boldsymbol{\xi}_t' S^{-1}\boldsymbol{\xi}_t - 2\boldsymbol{\xi}_t'\tilde{\boldsymbol{m}}_{tI_t}\}\right)\times\exp\left(-\frac{(\tau - I_t)y_t}{\lambda}\right), \quad t = 2, ..., n-1, \tag{41}$$

where

$$S = (T'(\sigma_\eta^2 Q)^{-1}T + (\sigma_\eta^2 Q)^{-1})^{-1}, \tag{42}$$

$$\tilde{\boldsymbol{m}}_{tI_t} = \left(T'(\sigma_\eta^2 Q)^{-1}\boldsymbol{\xi}_{t+1} + (\sigma_\eta^2 Q)^{-1}T\boldsymbol{\xi}_{t-1} + \frac{(\tau - I_t)}{\lambda}Z'\right). \tag{43}$$

Let

$$\tilde{\boldsymbol{m}}_{tI_t} = \begin{pmatrix} m_{tI_t} \\ \tilde{\boldsymbol{m}}_t \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{pmatrix},$$

where m_{tI_t}, S^{11} are scalars, $\tilde{\boldsymbol{m}}_t, (S^{12})', S^{21}$ are $(m-1)\times 1$ vectors and S^{22} is an $(m-1)\times(m-1)$ matrix. We generate $(\boldsymbol{\xi}_t, I_t)$ as follows:

a. $I_t|\tilde{\boldsymbol{\xi}}_t, \{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda.$

b. $\xi_t|I_t, \tilde{\boldsymbol{\xi}}_t, \{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda.$

c. $\tilde{\boldsymbol{\xi}}_t|\xi_t, \{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda.$

Noting that

$$f(\xi_t, I_t|\tilde{\boldsymbol{\xi}}_t, \{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda) \propto \exp\Big(-\frac{S^{11}}{2}\big\{\xi_t - (S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{tI_t})\big\}^2\Big) \\ \times \exp\Big(\frac{1}{2}(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{tI_t})^2\Big) \times \exp\Big(-\frac{(\tau - I_t)y_t}{\lambda}\Big), \tag{44}$$

define

$$g(0) := \Phi\Big(\frac{y_t - (S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t0})}{(S^{11})^{-\frac{1}{2}}}\Big) \times \exp\Big(\frac{1}{2}(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t0})^2 - \frac{\tau y_t}{\lambda}\Big), \tag{45}$$

$$g(1) := \Big(1-\Phi\Big(\frac{y_t - (S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t1})}{(S^{11})^{-\frac{1}{2}}}\Big)\Big)\times\exp\Big(\frac{1}{2}(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t+m_{t1})^2-\frac{(\tau-1)y_t}{\lambda}\Big), \tag{46}$$

where Φ is a cumulative normal distribution function.

Generation of I_t. Generate $I_t|\cdot \sim$ Bernoulli(p) with $p = g(1)/(g(0)+g(1))$[8] .

Generation of ξ_t. Given I_t, generate

$$\xi_t|\cdot \sim \begin{cases} \mathrm{TN}_{(-\infty,y_t]}((S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t0}), (S^{11})^{-1}) & \text{if } I_t = 0, \\ \mathrm{TN}_{(y_t,\infty)}((S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t1}), (S^{11})^{-1}) & \text{if } I_t = 1. \end{cases} \tag{49}$$

Generation of $\tilde{\boldsymbol{\xi}}_t$. Generate

$$\tilde{\boldsymbol{\xi}}_t|\xi_t, \{\boldsymbol{\xi}_{-t}\}, \{y_t\}_{t=1}^n, \sigma_\eta^2, \lambda \sim \mathrm{N}((S^{22})^{-1}(-S^{21}\xi_t + \tilde{\boldsymbol{m}}_t), (S^{22})^{-1}). \tag{50}$$

---

[8] When $\zeta_0 = [y_t - (S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t0})]/(S^{11})^{-\frac{1}{2}} \ll 0$, $\Phi(\zeta_0)$ may be almost 0 and $\exp\Big(\frac{1}{2}(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t0})^2 - \frac{\tau y_t}{\lambda}\Big)$ becomes huge. In this case, the computed value of g is inaccurate and we need to approximate the values of g using a partial fractional expansion (Stuart and Ord (1994)),

$$\Phi(\zeta_0) = \frac{1}{\sqrt{2\pi}}(-\zeta_0^{-1} + \zeta_0^{-3} - 3\zeta_0^{-5} + 15\zeta_0^{-7} - ...)\exp\Big(-\frac{\zeta_0^2}{2}\Big), \tag{47}$$

and obtain

$$g_t(0) \approx \frac{1}{\sqrt{2\pi}}\exp\Big(\log(-\zeta_0^{-1} + \zeta_0^{-3} - 3\zeta_0^{-5} + 15\zeta_0^{-7}) - \frac{\zeta_0^2}{2} + \frac{1}{2}(S^{11})^{-1}(-S^{12}\tilde{\boldsymbol{\xi}}_t + m_{t0})^2 - \frac{\tau}{\lambda}y_t\Big). \tag{48}$$

The approximation error is

$$\text{error} < \frac{105}{\sqrt{2\pi}}\zeta_0^{-9}\exp\Big(-\frac{\zeta_0^2}{2}\Big) < 6.34\times 10^{-14}$$

when $\zeta_0 \le -6$.

A.2 CAViaR model

As a benchmark for the model comparison, we consider the following asymmetric slope CAViaR model discussed in Engle and Manganelli (2004):

$$\xi_{t+1} = \beta_1 + \beta_2 \xi_t + \beta_3 y_t^+ + \beta_4 y_t^-, \tag{51}$$

where $y_t^+ = \max(y_t, 0)$ and $y_t^- = -\min(y_t, 0)$ to model asymmetry of the dynamics of the quantile. We assume the following prior distributions:

$$\lambda \sim \mathrm{IG}(0.1, 0.1), \;\; \boldsymbol{\beta} \sim \mathrm{TN}_{(0 \leq \beta_2 < 1)}(\mathbf{0}_4, 100 E_4), \;\; \xi_1 \sim \mathrm{N}(0, 100). \tag{52}$$

For $\tau = 0.1$ and for $\tau = 0.9$, respectively, we generate 240,000 MCMC draws after discarding 10,000 draws as the burn-in period.

A.3 Particle filter

Let $f(\boldsymbol{\xi}_t | Y_t, \boldsymbol{\theta})$ denote the density function of $\boldsymbol{\xi}_t$ given $(Y_t, \boldsymbol{\theta})$ where $Y_t = \{y_1, ..., y_t\}$, and let $\hat{f}(\xi_t | Y_t, \boldsymbol{\theta})$ denote the discrete approximation to $f(\boldsymbol{\xi}_t | Y_t, \boldsymbol{\theta})$.

We draw M samples from the conditional joint distribution of $(\boldsymbol{\xi}_{t+1}, \boldsymbol{\xi}_t, v_t)$ given $(Y_{t+1}, \boldsymbol{\theta})$ with the density

$$f(\boldsymbol{\xi}_{t+1}, \boldsymbol{\xi}_t, v_t | Y_{t+1}, \boldsymbol{\theta}) \propto f(y_{t+1} | \boldsymbol{\xi}_{t+1}, \boldsymbol{\theta}) f(\boldsymbol{\xi}_{t+1} | y_t, \boldsymbol{\xi}_t, v_t, \boldsymbol{\theta}) f(v_t | \boldsymbol{\theta}) f(\boldsymbol{\xi}_t | Y_t, \boldsymbol{\theta}), \tag{53}$$

where

$$f(y_t | \boldsymbol{\xi}_t, \boldsymbol{\theta}) = \int f(y_t | \boldsymbol{\xi}_t, v_t, \boldsymbol{\theta}) f(v_t | \boldsymbol{\theta}) dv_t, \;\; f(\boldsymbol{\xi}_{t+1} | y_t, \boldsymbol{\xi}_t, v_t, \boldsymbol{\theta}) = f(y_t, \boldsymbol{\xi}_{t+1} | \boldsymbol{\xi}_t, v_t, \boldsymbol{\theta}) / f(y_t | \boldsymbol{\xi}_t, v_t, \boldsymbol{\theta}). \tag{54}$$

We implement the particle filter:

1. (a) Generate

$$\boldsymbol{\xi}_1^{(i)} \sim \mathrm{N}(\boldsymbol{m}_1, V_1), \; i = 1, ..., M,$$

where $\boldsymbol{m}_1, V_1$ are some constant vector and some constant positive-definite matrix (we adopt the posterior mean and the covariance matrix of $\boldsymbol{\xi}_1$), respectively.

(b) We calculate

$$w_1^{(i)} := \frac{f(y_1 | \boldsymbol{\xi}_1^{(i)}, \boldsymbol{\theta}) f(\boldsymbol{\xi}_1^{(i)})}{g(\boldsymbol{\xi}_1^{(i)})},$$

where $g(\cdot)$ is a normal density with mean $\boldsymbol{m}_1$ and covariance matrix V_1 and set $\bar{w}_1 = \frac{1}{M}\sum_{i=1}^{M} w^{(i)}$, $\pi_1^{(i)} := \hat{f}(\boldsymbol{\xi}_1^{(i)}|y_1, \boldsymbol{\theta}^*) = w^{(i)}/\sum_{j=1}^{M} w^{(j)}$.

2. For $t = 1, ..., n-1$, we generate $(\boldsymbol{\xi}_{t+1}^{(i)}, \boldsymbol{\xi}_t^{(i)}), i = 1, ..., M$:

   (a) Generate $\boldsymbol{\xi}_t^{(i)} \sim \hat{f}(\boldsymbol{\xi}_t^{(i)}|Y_t, \boldsymbol{\theta})$ and $v_t^{(i)} \sim f(v_t^{(i)}|\boldsymbol{\theta})$.

   (b) Generate $\boldsymbol{\xi}_{t+1}^{(i)} \sim f(\boldsymbol{\xi}_{t+1}^{(i)}|y_t, \boldsymbol{\xi}_t^{(i)}, v_t^{(i)}, \boldsymbol{\theta})$.

   (c) Compute

$$\begin{aligned} w_{t+1}^{(i)} &:= \frac{f(y_{t+1}|\boldsymbol{\xi}_{t+1}^{(i)}, \boldsymbol{\theta}) f(\boldsymbol{\xi}_{t+1}^{(i)}|y_t, \boldsymbol{\xi}_t^{(i)}, v_t^{(i)}, \boldsymbol{\theta}) f(v_t^{(i)}|\boldsymbol{\theta}) \hat{f}(\boldsymbol{\xi}_t^{(i)}|Y_t, \boldsymbol{\theta})}{f(\boldsymbol{\xi}_{t+1}^{(i)}|y_t, \boldsymbol{\xi}_t^{(i)}, v_t^{(i)}, \boldsymbol{\theta}) f(v_t^{(i)}|\boldsymbol{\theta}) \hat{f}(\boldsymbol{\xi}_t^{(i)}|Y_t, \boldsymbol{\theta})} \\ &= f(y_{t+1}|\boldsymbol{\xi}_{t+1}^{(i)}, \boldsymbol{\theta}), \end{aligned}$$

   and set $\bar{w}_{t+1} = \frac{1}{M}\sum_{i=1}^{M} w^{(i)} \to f(y_{t+1}|Y_t, \boldsymbol{\theta}^*)$, $\hat{f}(\boldsymbol{\xi}_{t+1}^{(i)}|Y_{t+1}, \boldsymbol{\theta}^*) = w^{(i)}/\sum_{j=1}^{M} w^{(j)} := \pi_{t+1}^{(i)}$.

3. We obtain

$$\sum_{t=0}^{n-1} \log \bar{w}_{t+1} \to \sum_{t=0}^{n-1} \log f(y_{t+1}|Y_t, \boldsymbol{\theta}^*) = \log f(Y_n|\boldsymbol{\theta}^*) \text{ as } M \to \infty.$$

References

Chen, M., Q. Shao, and J. G. Ibrahim (2000). *Monte Carlo Methods in Bayesian Computation*. Springer-Verlag.

Dagpunar, J. S. (1989). An easily implemented generalized inverse Gaussian generator. *Communications in Statistics - Simulation and Computation 18*, 703–710.

de Jong, P. and N. Shephard (1995). The simulation smoother for time series models. *Biometrika 82*, 339–350.

De Rossi, G. and A. Harvey (2009). Quantiles, expectiles and splines. *Journal of Econometrics 152*, 179–185.

Doornik, J. A. (2007). *An Object-oriented Martrix Programming Language: Ox 5*. Timberlake Consultants Press.

Doucet, A., N. de Freitas, and N. Gordon (Eds.) (2001). *Sequential Monte Carlo Methods in Practice*. Springer.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford University Press.

Durbin, J. and S. J. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika 89*, 603–615.

Engle, R. F. (Ed.) (1995). *ARCH: Selected Readings*. Oxford University Press.

Engle, R. F. and S. Manganelli (2004). CAViaR: Conditional Autoregressive Value at Risk by regression quantiles. *Journal of Business & Economic Statistics 22*, 367–381.

Gerlach, R. H., C. W. S. Chen, and N. Y. C. Chan (2011). Bayesian time-varying quantile forecasting for Value-at-Risk in financial markets. *Journal of Business & Economic Statistics 29*, 481–492.

Gourieroux, C. and J. Jasiak (2008). Dynamic quantile models. *Journal of Econometrics 147*, 198–205.

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

Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.

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

Koenker, R. and J. A. F. Machado (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association 94*, 1296–1310.

Koenker, R. and Z. Xiao (2006). Quantile autoregression. *Journal of the American Statistical Association 101*, 980–990.

Kohn, R. and C. F. Ansley (1987). A new algorithm for spline smoothing based on smoothing a stochastic process. *SIAM Journal on Scientific and Statistical Computing 8*, 33–48.

Kotz, S., T. J. Kozubowski, and K. Podgórski (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Birkhäuser.

Kozumi, H. and G. Kobayashi (2011). Gibbs sampling methods for Bayesian quantile regression. *Journal of Statistical Computation and Simulation 81*, 1565–1578.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives 3*, 73–84.

Morf, M. and T. Kailath (1975). Square-root algorithms for least-squares estimation. *IEEE Transactions on Automatic Control 4*, 487–497.

Omori, Y., S. Chib, N. Shephard, and J. Nakajima (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics 140*, 425–449.

Shephard, N. (Ed.) (2005). *Stochastic Volatility: Selected Readings*. Oxford University Press.

Shephard, N. and M. K. Pitt (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika 84*, 653–667.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Ser. B 64*, 583–639.

Stuart, A. and K. Ord (1994). *Kendall's Advanced Theory of Statistics* (6 ed.), Volume 1. Arnold.

Tsionas, E. G. (2003). Bayesian quantile inference. *Journal of Statistical Computation and Simulation 73*, 659–674.

Watanabe, T. and Y. Omori (2004). A multi-move sampler for estimating non-Gaussian time series models: Comments on Shephard & Pitt (1997). *Biometrika 91*, 246–248.

Wecker, W. E. and C. F. Ansley (1983). The signal extraction approach to nonlinear regression and spline smoothing. *Journal of the American Statistical Association 78*, 81–89.

Yu, K. and R. A. Moyeed (2001). Bayesian quantile regression. *Statistics & Probability Letters 54*, 437–447.

Yue, Y. R. and H. Rue (2011). Bayesian inference for additive mixed quantile regression models. *Computational Statistics & Data Analysis 55*, 84–96.