Time-Varying Parameters Bayesian Quantile Regression*

Shiqi Ye

Xiamen University

June 2022

Abstract

Recent years, due to some important global events such as the COVID-19 and the Russia-Ukraine conflict, the global economic growth has been stuck and a steep downward trend is appear, indicating an forthcoming economic contraction. Therefore, the measurement of tail risks of economic growth or inflation has become an important research problem. In this paper, I first follow Korobilis (2017) and develop the estimation procedure for Bayesian quantile regression. Then I draw on idea in Korobilis *et al.* (2021) and construct a time-varying parameters Bayesian quantile regression. The real data exercise shows that the proposed models have similar performance with the frequentist 's approach. Finally, I use these models to investigate a AR(2) specification for CPI in different quantiles. The results show two important findings. First, the coefficients of lag terms of CPI show significant time-varying effects. Second, compared with the median or 0.95 quantile, in the case of extreme left tail (0.05 quantile), the lag items of CPI can better reflect the downward risk, which was most obvious during the 2008 global financial crisis.

*I would like to thank professor Weixuan Zhu for many helpful comments.

1 The quantile regression

Consider a weakly stationary time series $\{y_t\}_{t=1}^{T}$, we start with the definition of quantile function. For a scalar random variable Y, a *quantile function* of Y is defined as a inverse of its distribution. The quantile function, similar with the distribution, is an alternative statistic property of Y. Mathematically, the quantile function of Y can be expressed as:

$$\mathcal{Q}_Y(p) = F_Y^{-1}(p) = \inf\{y : F_Y(y) \geq p\},$$

where $p \in [0, 1]$ denotes the quantile of interest and $F_Y(y) = p(Y \leq y)$. Similarly, the *conditional quantile function* of Y given random variable X can be expressed as:

$$\mathcal{Q}_{Y|X}(p|X) = F_{Y|X}^{-1}(p|X) = \inf\{y : F_{Y|X}(y|X) \geq p\}, \tag{1.1}$$

where $F_{Y|X}(y) = p(Y \leq y|X)$. The conditional quantile function can totally reflect the relationship between Y and X. Based on Equ.(1.1), let us consider a linear model below:

$$y_t = \boldsymbol{x}_t' \boldsymbol{\beta} + \epsilon_t,$$

where $\boldsymbol{x}_t$ denotes $N \times 1$ vector of exogenous variables and own lags, $\boldsymbol{\beta}$ is a $N \times 1$ vector of coefficients and ϵ_t is $i.i.d$ with mean zero and constant variance. A regression of the above model can be expressed as:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \sum_{t=1}^{T} \rho\left(y_t - \boldsymbol{x}_t'\boldsymbol{\beta}\right), \tag{1.2}$$

where $\rho(\cdot)$ denotes a loss function or criterion function. Different $\rho(\cdot)$ will result in different estimator $\hat{\boldsymbol{\beta}}$, which has completely different interpretations.

For example, if we let $\rho(\mu) = \mu^2$, then we obtain the OLS estimator $\hat{\boldsymbol{\beta}}_{\text{ols}}$, which is a consistent estimator of $\boldsymbol{\beta}^*_{\text{ols}}$ with $\boldsymbol{x}_t'\boldsymbol{\beta}^*_{\text{ols}} = E(Y|X)$. If we let $\rho(\mu) = |\mu|$, then we obtain the Least Absolute Deviation (LAD) estimator $\hat{\boldsymbol{\beta}}_{\text{lad}}$, which is a consistent estimator of $\boldsymbol{\beta}^*_{\text{lad}}$ with $\boldsymbol{x}_t'\boldsymbol{\beta}^*_{\text{lad}} = \text{Median}(Y|X)$.

Koenker and Bassett Jr (1978), alternatively, consider a check function by letting

$$\rho(\mu) = \mu(p - I(\mu < 0)),$$

where $p \in (0,1)$ and $I(\cdot)$ denotes an indicator function. Note that

$$\rho(\mu) = (1-p)I[\mu < 0]|\mu| + pI[\mu \geq 0]|\mu|,$$

which implies an asymmetric and weighted sum of absolute errors. Koenker and Bassett Jr (1978) show that the estimator $\hat{\boldsymbol{\beta}}_{\mathrm{p}}$ is a consistent estimator of $\boldsymbol{\beta}^*_{\mathrm{p}}$. The linear model that solve the above minimizing problem in Equ.(1.2) under the above check function is the so-called quantile regression. The quantile regression aims to portray the conditional probability distribution of random variable Y given X at its quantile p, and has widely been used in analyzing the impact of tail-risk event on the economic indices.

2 Bayesian quantile regression

2.1 Model specification

Yu and Moyeed (2001) show that the minimizing problem of quantile regression, as discussed in the previous section, is equivalent to a maximization problem of likelihood function based on error term that follows independent asymmetric Laplace distribution. They implement a random walk Metropolis–Hastings (MH) algorithm with a Gaussian proposal density to estimate the posterior of $\boldsymbol{\beta}_p$ at quantile p. The idea is straightforward and convenient. however, a main drawback of the MH algorithm for quantile regression is that when p varies, the MH algorithm tends to have different accept rates, and the tuning parameters need to be adjusted for each p to ensure a good performance.

To address this issue, Kozumi and Kobayashi (2011) propose a convenient and efficient algorithm by first demonstrating that the random error in the quantile regression is actually equivalent with the following form:

$$\epsilon_t = \theta z_{t,p} + \tau\sqrt{z_{t,p}}\mu_t, \tag{2.1}$$

where $z_{t,p} \sim \mathrm{Exp}(1)$ denotes an exponential distribution with the scale parameter equals to 1, $\mu_t \sim N(0,1)$ denotes standard normal distribution, $\theta = (1-2p)/(p * (1-p))$ and $\tau^2 = 2/p(1-p)$.

Then, the Bayesian quantile regression (BQR) is equivalent to the following representation:

$$y_t = \boldsymbol{x}_t' \boldsymbol{\beta}_p + \theta z_{t,p} + \tau \sqrt{z_{t,p}} \mu_t. \tag{2.2}$$

An appealing property of the above representation is that the conditional distribution of y_t given $\boldsymbol{\beta}_p$ and $z_{t,p}$ is normal, and we have:

$$p(\boldsymbol{y}|\boldsymbol{\beta}_p, \boldsymbol{z}_p) \propto \left(\prod_{t=1}^{T} z_{t,p}^{-\frac{1}{2}} \right) \times \exp \left\{ -\frac{1}{2} \sum_{t=1}^{T} \frac{(y_t - \boldsymbol{x}_t' \boldsymbol{\beta}_p - \theta z_{t,p})^2}{\tau^2 z_{t,p}} \right\}, \tag{2.3}$$

where $\boldsymbol{y} = (y_1, \ldots, y_T)'$ and $\boldsymbol{z}_p = (z_1, \ldots, z_{t,p})'$.

2.2 Prior settings and posterior inference

Given the likelihood function of $\boldsymbol{y}$, the next step is to specify the prior distribution for $\boldsymbol{\beta}_p$ and $z_{t,p}$. I access the stochastic search variable selection (SSVS) prior for $\boldsymbol{\beta}_p$ considered in Korobilis (2017) since I'm pretty interested with his research and its potential applications. Specifically, the prior has the form:

$$\beta_{i,p}|\gamma_{i,p}, \delta_{i,p} \sim (1 - \gamma_{i,p}) N(0, \underline{c} \times \delta_{i,p}^2) + \gamma_{i,p} N(0, \delta_{i,p}^2), \tag{2.4}$$

$$\delta_{i,p}^{-2} \sim Gamma(\underline{a}_1, \underline{a}_2), \tag{2.5}$$

$$\gamma_{i,p}|\pi_0 \sim Bernoulli(\pi_0), \tag{2.6}$$

$$\pi_0 \sim Beta(\underline{b}_1, \underline{b}_2). \tag{2.7}$$

Where $(\underline{a}_1, \underline{a}_2, \underline{b}_1, \underline{b}_2)$ are pre-determined hyperparameters, and $\underline{c}$ denotes a fixed constant that is very close to zero.

To implement the Gibbs sampler, we need to first derive the conditional posterior distribution of $\boldsymbol{\beta}_p, \delta_{i,p}^2, \gamma_{i,p}$ and π_0 for $i = 1, \ldots, N$. First, let $\boldsymbol{\gamma}_p = (\gamma_{1,p}, \ldots, \gamma_{k,p})'$

and $\boldsymbol{\delta}_p = (\delta_{1,p}, \ldots, \delta_{k,p})'$, $\boldsymbol{x} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T\}$ the full posterior distribution becomes:

$$\begin{aligned}
p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}, \pi_0 | \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}_p) &\propto p(\boldsymbol{y}|\boldsymbol{\beta}_p, \boldsymbol{z}_p) p(\boldsymbol{\beta}_p|\boldsymbol{\gamma}, \boldsymbol{\delta}) p(\boldsymbol{\gamma}|\pi_0) p(\boldsymbol{\delta}) p(\pi_0) p(\boldsymbol{z}_p) \\
&\propto \left(\prod_{t=1}^{T} z_{t,p}^{-\frac{1}{2}} \right) \times \exp\left\{ -\frac{1}{2} \sum_{t=1}^{T} \frac{(y_t - \boldsymbol{x}_t' \boldsymbol{\beta}_p - \theta z_{t,p})^2}{\tau^2 z_{t,p}} \right\} \\
&\quad \times \left(\prod_{i=1}^{N} (\delta_i^{-2})^{\frac{1}{2}} \underline{c}^{-\frac{1-\gamma_i}{2}} \right) \times \exp\left\{ -\frac{1}{2} \delta_{i,p}^{-2} \sum_{i=1}^{N} \left(\gamma_i \beta_{i,p}^2 + (1 - \gamma_{i,p}) \frac{\beta_{i,p}^2}{\underline{c}} \right) \right\} \\
&\quad \times \prod_{i=1}^{N} \pi_0^{\gamma_i} (1 - \pi_0)^{1-\gamma_i} \times \prod_{i=1}^{N} (\delta_{i,p}^{-2})^{\underline{a}_1 - 1} \exp\{-\underline{a}_2 \delta_{i,p}^{-2}\} \\
&\quad \times \pi_0^{\underline{b}_1 - 1} (1 - \pi_0)^{\underline{b}_2 - 1} \times \exp\{-\sum_{t=1}^{T} z_{t,p}\}
\end{aligned}$$

Then, the conditional posterior inference can be obtained as follows:

1. The posterior inference for $\boldsymbol{\beta}_p$

Let us first look at $\boldsymbol{\beta}_p$, since $\gamma_{i,p}$ follows a discrete bernoulli prior distribution, which is either equal to 1 or 0, let $\boldsymbol{V}_p^{\beta}$ be a diagonal matrix with its kth element and define:

$$(\boldsymbol{V}_p^{\beta})_k = \begin{cases} \underline{c} \times \delta_{i,p}^2 & \text{if } \gamma_{k,p} = 0; \\ \delta_{i,p}^2 & \text{if } \gamma_{k,p} = 1, \end{cases} \tag{2.8}$$

we have:

$$\boldsymbol{\beta}_p | \boldsymbol{\gamma}_p, \boldsymbol{\delta}_p \ \sim \ \boldsymbol{N}\left(\boldsymbol{0}, \boldsymbol{V}_p^{\beta}\right).$$

Note that we also have:

$$\boldsymbol{y}|\boldsymbol{\beta}_p, \boldsymbol{x}, \boldsymbol{z}_p \sim \boldsymbol{N}\left(\theta \boldsymbol{z}_p + \boldsymbol{X}\boldsymbol{\beta}_p, \boldsymbol{V}_p^{y}\right),$$

where $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)'$ with $\boldsymbol{V}_p^y = \tau^2 \text{diag}\{z_1, \ldots, z_{t,p}\}$, then we have:

$$\begin{aligned}
p(\boldsymbol{\beta}_p | \boldsymbol{\gamma}_p, \boldsymbol{\delta}_p, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}_p) &\propto \exp\left\{ -\frac{1}{2} (\boldsymbol{y} - \theta \boldsymbol{z}_p - \boldsymbol{X}\boldsymbol{\beta}_p)' \left(\boldsymbol{V}_p^y\right)^{-1} (\boldsymbol{y} - \theta \boldsymbol{z}_p - \boldsymbol{X}\boldsymbol{\beta}_p) - \frac{1}{2} \boldsymbol{\beta}_p' (\boldsymbol{V}_p^{\beta})^{-1} \boldsymbol{\beta}_p \right\} \\
&\propto \exp\left\{ -\frac{1}{2} \left(-2\boldsymbol{\beta}_p' (\boldsymbol{V}_p^y)^{-1} (\boldsymbol{y} - \theta \boldsymbol{z}_p) + \boldsymbol{\beta}_p' \boldsymbol{X} (\boldsymbol{V}_p^y)^{-1} \boldsymbol{X}' \boldsymbol{\beta}_p + \boldsymbol{\beta}_p' \boldsymbol{X} (\boldsymbol{V}_p^{\beta})^{-1} \boldsymbol{\beta}_p \right) \right\} \\
&\propto \exp\left\{ -\frac{1}{2} \left(\boldsymbol{\beta}_p - \overline{\boldsymbol{\beta}}_p\right)' (\overline{\boldsymbol{V}}_p^{\beta})^{-1} \left(\boldsymbol{\beta}_p - \overline{\boldsymbol{\beta}}_p\right) \right\},
\end{aligned}$$

so that

$$\boldsymbol{\beta}_p|\boldsymbol{\gamma}_p, \boldsymbol{\delta}_p, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}_p \ \sim \ \boldsymbol{N}(\overline{\boldsymbol{\beta}}_p, \overline{\boldsymbol{V}}_p^{\beta}), \tag{2.9}$$

with:

$$\overline{\boldsymbol{V}}_p^{\beta} = \left(\boldsymbol{X}'(\boldsymbol{V}_p^y)^{-1}\boldsymbol{X} + \boldsymbol{V}_p^{\beta}\right)^{-1} = \left(\sum_{t=1}^{T} \frac{\boldsymbol{x}_t\boldsymbol{x}_t'}{\tau^2 z_{t,p}} + \boldsymbol{V}_p^{\beta}\right)^{-1},$$

and

$$\overline{\boldsymbol{\beta}}_p = \overline{\boldsymbol{V}}_p^{\beta}\boldsymbol{X}'(\boldsymbol{V}_p^y)^{-1}\left(\boldsymbol{y} - \theta\boldsymbol{z}_p\right) = \overline{\boldsymbol{V}}_p^{\beta}\sum_{t=1}^{T} \frac{\boldsymbol{x}_t(y_t - \theta z_{t,p})}{\tau^2 z_{t,p}}.$$

2. The posterior inference for $\delta_{i,p}^{-2}$

When $\gamma_i = 1$, we have:

$$p(\delta_i^{-2}|\beta_{i,p}, \gamma_i = 1) \propto (\delta_i^{-2})^{\underline{a}_1 + \frac{1}{2} - 1} \exp\left\{-\left(\frac{\beta_{i,p}^2}{2} + \underline{a}_2\right)\delta_i^{-2}\right\}.$$

When $\gamma_i = 0$, the SSVS indicates that $\beta_{i,p} = 0$, and we have:

$$p(\delta_i^{-2}|\beta_{i,p}, \gamma_i = 0) \propto (\delta_i^{-2})^{\underline{a}_1 + \frac{1}{2} - 1} \exp\left\{-\underline{a}_2\delta_i^{-2}\right\}.$$

Combine the above results together, we have:

$$\delta_{i,p}^{-2} \ \sim \ Gamma(\overline{a}_1, \overline{a}_2), \tag{2.10}$$

with

$$\overline{a}_1 = \underline{a}_1 + \frac{1}{2}, \qquad \text{and} \qquad \overline{a}_2 = \frac{\beta_{i,p}^2}{2} + \underline{a}_2.$$

3. The posterior inference for $\gamma_{i,p}$

Define $\boldsymbol{\gamma}_{-i,p}$ be the collection of $\gamma_{j,p}$s with $j \neq i$, then we have:

$$p(\gamma_{i,p} = 1|\boldsymbol{\gamma}_{-i,p}, \boldsymbol{\beta}_p, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}_p) \propto (\delta_{i,p}^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\delta_{i,p}^2}\beta_{i,p}^2\right\}\pi_0,$$

and

$$p(\gamma_{i,p}=0|\boldsymbol{\gamma}_{-i,p},\boldsymbol{\beta}_p,\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}_p) \propto (\underline{c}\delta_{i,p}^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\underline{c}\delta_{i,p}^2}\beta_{i,p}^2\right\}(1-\pi_0),$$

so we have:

$$\gamma_{i,p}|\boldsymbol{\gamma}_{-i,p},\beta_{i,p},\pi_0 \ \sim \ Bernoulli(\overline{\pi}), \tag{2.11}$$

with

$$\overline{\pi} = \frac{\pi_0\phi(\beta_{i,p}|0,\delta_{i,p}^2)}{\pi_0\phi(\beta_{i,p}|0,\delta_{i,p}^2)+(1-\pi_0)\phi(\beta_{i,p}|0,\underline{c}\delta_{i,p}^2)},$$

where $\phi(\beta_{i,p}|0,\delta_{i,p}^2)$ denotes the normal probability density distribution evaluated at $\beta_{i,p}$ with zero mean and variance $\delta_{i,p}^2$.

4. The posterior inference for π_0

Given $\pi_0 \ \sim \ Beta(\underline{b}_1,\underline{b}_2)$, the conditional posterior distribution of β can be obtained by:

$$p(\pi_0|\boldsymbol{\gamma}_p) \propto \pi_0^{\underline{b}_1+\sum_{i=1}^N \gamma_{i,p}-1}(1-\pi_0)^{\underline{b}_2+N-\sum_{i=1}^N(\gamma_{i,p})-1}.$$

Thus, we can obtain the conditional posterior distribution of π_0:

$$\pi_0|\boldsymbol{\gamma}_p \ \sim \ Beta(\overline{b}_1,\overline{b}_2), \tag{2.12}$$

where

$$\overline{b}_1 = \underline{b}_1+\sum_{i=1}^N\gamma_{i,p}, \qquad \text{and} \qquad \overline{b}_2 = \underline{b}_2+N-\sum_{i=1}^N\gamma_{i,p}.$$

5. The posterior inference for $z_{t,p}$

Conditional on other parameters and the data, we have:

$$\begin{aligned} p(z_{t,p}|\boldsymbol{x},\boldsymbol{y},\boldsymbol{\beta}_p) &\propto z_{t,p}^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2}\frac{(y_t-\boldsymbol{x}_t'\boldsymbol{\beta}_p-\theta z_{t,p})^2}{\tau^2 z_{t,p}} - z_{t,p}\right\} \\ &\propto z_{t,p}^{\frac{1}{2}-1} \times \exp\left\{-\frac{1}{2}\left(\frac{(y_t-\boldsymbol{x}_t'\boldsymbol{\beta}_p)^2}{\tau^2}z_{t,p}^{-1}+\frac{2\tau^2+\theta^2}{\tau^2}z_{t,p}\right)\right\}. \end{aligned}$$

Then we can conclude that:

$$z_{t,p} \;\sim\; GIG(\frac{1}{2}, \overline{\kappa}_1, \overline{\kappa}_2), \tag{2.13}$$

with

$$\overline{\kappa}_1 = \frac{|y_t - \boldsymbol{x}_t'\boldsymbol{\beta}_p|}{\tau}, \qquad \text{and} \qquad \overline{\kappa}_2 = \frac{\sqrt{2\tau^2 + \theta^2}}{\tau},$$

where GIG represents the generalized inverse gaussian distribution $GIG(\nu, a, b)$, the pdf of which has the form:

$$f(x) = \frac{(a/b)^\nu}{2K_\nu(\sqrt{ab})} x^{\nu-1} \exp\left\{-\frac{1}{2}(a^2x^{-1} + b^2x)\right\}.$$

It is hard to directly sample from the GIG distribution. However, since $\nu = \frac{1}{2}$ in our case, there exists a tricky way to sample from GIG distribution. Let $X \sim GIG(\frac{1}{2}, \overline{\kappa}_1, \overline{\kappa}_2)$ and that $Z = 1/X$, then we have the pdf of Z is proportional to:

$$p(z) \propto z^{\frac{1}{2}} \times z^{-2} \exp\left\{-\frac{1}{2}(\overline{\kappa}_2 z^{-1} + \overline{\kappa}_1 z)\right\}.$$

We have $Z \sim GIG(-\frac{1}{2}, \overline{\kappa}_2, \overline{\kappa}_1)$, which is equivalent to a inverse Gaussian distribution $IG(\mu, \gamma)$ with $\mu = \overline{\kappa}_2/\overline{\kappa}_1$ and $\lambda = \overline{\kappa}_2^2$.

That is, we first sample Z from the $IG(\mu, \lambda)$ with

$$\mu = \frac{\sqrt{2\tau^2 + \theta^2}}{|y_t - \boldsymbol{x}_t'\boldsymbol{\beta}_p|} \qquad \text{and} \qquad \lambda = \frac{2\tau^2 + \theta^2}{\tau^2}$$

2.3 The MCMC algorithm

Given the conditional posterior distribution, it is natural to implement Gibbs sampler to sample the parameters from the posterior distribution. Although the key steps of Gibbs sampler has been discussed in the Appendix B of Korobilis (2017), there are some typos and we correct it in this paper. The Gibbs sampler is implemented by sequentially sampling from the conditional posterior distributions:

1. Sample $\boldsymbol{\beta}_p$ conditional on other parameters and the data:

$$\boldsymbol{\beta}_p | \boldsymbol{\gamma}_p, \boldsymbol{\delta}_p, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}_p \ \sim \ \boldsymbol{N}(\overline{\boldsymbol{\beta}}_p, \overline{\boldsymbol{V}}_p^{\beta}),$$

where $\overline{\boldsymbol{V}}_p^{\beta} = \left(\sum_{t=1}^{T} \frac{\boldsymbol{x}_t \boldsymbol{x}_t'}{\tau^2 z_{t,p}} + (\boldsymbol{V}_p^{\beta})^{-1} \right)^{-1}$, $\overline{\boldsymbol{\beta}}_p = \overline{\boldsymbol{V}}_p^{\beta} \sum_{t=1}^{T} \frac{\boldsymbol{x}_t (y_t - \theta z_{t,p})}{\tau^2 z_{t,p}}$ and $\boldsymbol{V}_p^{\beta}$ be a diagonal matrix which is defined in Equ.(2.8).

2. Sample $\delta_{i,p}^2$ for $i = 1, \ldots, N$, conditional on other parameters and the data:

$$\delta_{i,p}^{-2} \ \sim \ Gamma(\overline{a}_1, \overline{a}_2),$$

where $\overline{a}_1 = \underline{a}_1 + \frac{1}{2}$ and $\overline{a}_2 = \frac{\beta_{i,p}^2}{2} + \underline{a}_2$.

3. Sample $\gamma_{i,p}$ for $i = 1, \ldots, N$, conditional on other parameters and the data:

$$\gamma_{i,p} | \boldsymbol{\gamma}_{-i,p}, \beta_{i,p}, \pi_0 \ \sim \ Bernoulli(\overline{\pi}),$$

where $\overline{\pi} = \frac{\pi_0 \phi(\beta_{i,p}|0, \delta_{i,p}^2)}{\pi_0 \phi(\beta_{i,p}|0, \delta_{i,p}^2) + (1-\pi_0)\phi(\beta_{i,p}|0, \underline{c}\delta_{i,p}^2)}$ and $\phi(\beta_{i,p}|0, \delta_{i,p}^2)$ denotes the normal probability density distribution evaluated at $\beta_{i,p}$ with zero mean and variance $\delta_{i,p}^2$.

4. Sample $z_{t,p}$ for $t = 1, \ldots, T$, conditional on other parameters and the data:

$$z_{t,p} \ \sim \ GIG(1, \overline{\kappa}_1, \overline{\kappa}_2),$$

where $\overline{\kappa}_1 = \frac{|y_t - \boldsymbol{x}_t' \boldsymbol{\beta}_p|}{\tau}$ and $\overline{\kappa}_2 = \frac{\sqrt{2+\theta^2}}{\tau}$.

5. Sample π_0 conditional on other parameters and the data:

$$\pi_0 | \boldsymbol{\gamma}_p \ \sim \ Beta(\overline{b}_1, \overline{b}_2),$$

where $\overline{b}_1 = \underline{b}_1 + \sum_{i=1}^{N} \gamma_{i,p}$ and $\overline{b}_2 = \underline{b}_2 + N - \sum_{i=1}^{N} \gamma_{i,p}$.

3 Time-varying Parameter BQR

Researchers always believe that the economic indicators have structural changes (see, for example, Stock and Watson (1996); Kim and Nelson (1999); Bernanke (2004); Justiniano and Primiceri (2008); Welch and Goyal (2008); Benati and Surico (2009); Rossi (2013); Hong *et al.* (2017)). In this case, some traditional constant coefficient models may not be suitable, including the Bayesian quantile regression. Korobilis *et al.* (2021) therefore extends the regression setting in section 2 by letting parameters to evolve through time.

3.1 Model specification

Lets consider a similar specification of Equ.(2.2), which is an extension of Khare and Hobert (2012):

$$y_t = \boldsymbol{x}_t'\boldsymbol{\beta}_{t,p} + \theta z_{t,p} + \sigma_p \tau \sqrt{z_{t,p}}\mu_t, \qquad \mu_t \sim N(0,1) \tag{3.1}$$

$$\boldsymbol{\beta}_{t,p} = \boldsymbol{\beta}_{t-1,p} + \boldsymbol{v}_{t,p}, \qquad \boldsymbol{v}_{t,p} \sim N(0, \boldsymbol{V}_p). \tag{3.2}$$

The difference between above equation with the model specification in the previous section lies in two parts. First, different from letting $z_{t,p} \sim \text{Exp}(1)$, here we consider a scaled $z_{t,p}$ (Kozumi and Kobayashi, 2011) by letting $z_{t,p} \sim \text{Exp}(\sigma_p^2)$. Second, the $\boldsymbol{\beta}_{t,p}$ is now following a random walk process, such a setting is standard in state space modeling and is convenient for Bayesian inference (Durbin and Koopman, 2012). Now let me introduce two reparameterizations of the above equations that are helpful to construct the Gibb's sampler:

Reparameterizaiton 1. If we conditional on $z_{t,p}$, then we have:

$$\begin{aligned}
y_t - \theta z_{t,p} &= \boldsymbol{x}_t'\boldsymbol{\beta}_{t,p} + \sigma_p \tau \sqrt{z_{t,p}}\mu_t, \\
\frac{y_t - \theta z_{t,p}}{\tau\sqrt{z_{t,p}}} &= \left(\frac{\boldsymbol{x}_t}{\tau\sqrt{z_{t,p}}}\right)' \boldsymbol{\beta}_{t,p} + \sigma_p \mu_t, \\
\tilde{y}_t &= \tilde{\boldsymbol{x}}_t'\boldsymbol{\beta}_{t,p} + \tilde{\epsilon}_t,
\end{aligned} \tag{3.3}$$

where $\tilde{\epsilon}_t \sim N(0, \sigma_p^2)$. Thus it is trivial to derive the conditional posterior distributions for $\boldsymbol{\beta}_{t,p}$ and σ_p^2 in this case.

Reparameterizaiton 2. Given the random walk specification of $\boldsymbol{\beta}_{t,p}$, Korobilis (2021) propose an efficient way by rewriting Equ.3.1:

$$\begin{aligned}
\tilde{y}_t &= \tilde{\boldsymbol{x}}_t'\boldsymbol{\beta}_{t,p} + \tilde{\epsilon}_t \\
&= \tilde{\boldsymbol{x}}_t'\Delta\boldsymbol{\beta}_{t,p} + \tilde{\boldsymbol{x}}_t'\boldsymbol{\beta}_{t-1,p} + \tilde{\epsilon}_t \\
&= \tilde{\boldsymbol{x}}_t'\Delta\boldsymbol{\beta}_{t,p} + \tilde{\boldsymbol{x}}_t'\Delta\boldsymbol{\beta}_{t-1,p} + \tilde{\boldsymbol{x}}_t'\boldsymbol{\beta}_{t-2,p} + \tilde{\epsilon}_t \\
&\cdots \\
&= \tilde{\boldsymbol{x}}_t'\Delta\boldsymbol{\beta}_{t,p} + \tilde{\boldsymbol{x}}_t'\Delta\boldsymbol{\beta}_{t-1,p} + \cdots + \tilde{\boldsymbol{x}}_t'\boldsymbol{\beta}_{1,p}\tilde{\epsilon}_t,
\end{aligned}$$

where $\Delta\boldsymbol{\beta}_{t,p} = \boldsymbol{\beta}_{t,p} - \boldsymbol{\beta}_{t-1,p}$. This is equivalent with:

$$\tilde{\boldsymbol{y}} = \mathcal{X}\boldsymbol{\beta}_p^{\Delta} + \tilde{\boldsymbol{\varepsilon}}, \tag{3.4}$$

Where $\tilde{\boldsymbol{y}} = (\tilde{y}_1, \ldots, \tilde{y}_T)'$, $\tilde{\boldsymbol{\varepsilon}} = (\tilde{\epsilon}_1, \ldots, \tilde{\epsilon}_T)$. $\mathcal{X}$ is a $T \times TN$ observation matrix and $\boldsymbol{\beta}_p^{\Delta}$ denotes a coefficient vector of dimension TN with:

$$\mathcal{X} = \begin{bmatrix} \tilde{\boldsymbol{x}}_1' & 0 & \cdots & 0 & 0 \\ \tilde{\boldsymbol{x}}_2' & \tilde{\boldsymbol{x}}_2' & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \tilde{\boldsymbol{x}}_{T-1}' & \tilde{\boldsymbol{x}}_{T-1}' & \tilde{\boldsymbol{x}}_{T-1}' & \tilde{\boldsymbol{x}}_{T-1}' & 0 \\ \tilde{\boldsymbol{x}}_T' & \tilde{\boldsymbol{x}}_T' & \tilde{\boldsymbol{x}}_T' & \tilde{\boldsymbol{x}}_T' & \tilde{\boldsymbol{x}}_T' \end{bmatrix}, \quad \boldsymbol{\beta}_p^{\Delta} = \begin{bmatrix} \boldsymbol{\beta}_{1,p} \\ \Delta\boldsymbol{\beta}_{2,p} \\ \cdots \\ \Delta\boldsymbol{\beta}_{t-1,p} \\ \Delta\boldsymbol{\beta}_{t,p} \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_{1,p} \\ \boldsymbol{v}_{2,p} \\ \cdots \\ \boldsymbol{v}_{T-1,p} \\ \boldsymbol{v}_{T,p} \end{bmatrix}$$

3.2 Prior settings

Given the model specification in Equ.(3.1), following Korobilis *et al.* (2021), we assume a horseshoe prior (Carvalho *et al.*, 2010) for $\boldsymbol{\beta}_p^{\Delta}$. The prior distribution for all the unknown parameters are:

$$\begin{aligned} \boldsymbol{\beta}_p^{\Delta} | \lambda_p^2, \{\psi_{i,p}^2\}_{i=1}^{TN} &\sim N(0, \boldsymbol{V}_p), \\ \boldsymbol{V}_{ii,p} &= \lambda_p^2 \psi_{i,p}^2, \quad i = 1, \ldots, TN, \\ \lambda_p &\sim Cauchy^+(0,1), \\ \psi_{i,p} &\sim Cauchy^+(0,1), \quad i = 1, \ldots, TN, \\ \sigma_p^2 &\sim \mathcal{IG}(\underline{\rho}_1, \underline{\rho}_2), \\ z_{t,p} &\sim exp(\sigma_p^2), \end{aligned}$$

where $\underline{\rho}_1, \underline{\rho}_2$ are two hyper-parameters that control the scale parameter of exponential distribution $z_{t,p}$.

According to Makalic and Schmidt (2015), the horseshoe prior for $\boldsymbol{\beta}_p^{\Delta}$ can be rewritten as a mixture of inverse Gamma distributions:

$$\begin{aligned} \boldsymbol{\beta}_p^{\Delta} | \lambda_p^2, \{\psi_{i,p}^2\}_{i=1}^{TN} &\sim N(0, \boldsymbol{V}_p), \\ \boldsymbol{V}_{ii,p} &= \lambda_p^2 \psi_{i,p}^2, \quad i = 1, \ldots, TN, \\ \lambda_p^2 | \xi_p &\sim \mathcal{IG}(1/2, 1/\xi_p), \\ \xi_p &\sim \mathcal{IG}(1/2, 1), \\ \psi_{i,p}^2 | \zeta_{i,p} &\sim \mathcal{IG}(1/2, 1/\zeta_{i,p}), \\ \zeta_{i,p} &\sim \mathcal{IG}(1/2, 1). \end{aligned}$$

3.3 Posterior inference

Given the piror settings, the next step is to inference the conditional posterior distributions for $\boldsymbol{\beta}_p^{\Delta}, \lambda_p^2, \xi_p, \psi_{i,p}^2, \zeta_{i,p}, z_{t,p}, \sigma_p^2$ for $i = 1, \ldots, TN$ given the data. Here we use $\bullet$ to denotes conditional on other parameters and data.

1. The posterior inference for $\boldsymbol{\beta}_p^{\Delta}$

The derivation of the conditional posterior distribution for $\boldsymbol{\beta}_p^{\Delta}$ is the same as in section 2.2 and has similar form with Equ.2.9, and the result has a little bit difference since in this case the $\boldsymbol{V}_p^y = \sigma_p^2 \tau^2 \text{diag}\{z_{1,p}, \ldots, z_{T,p}\}$. And we have:

$$\boldsymbol{\beta}_p^{\Delta}|\bullet \;\sim\; \boldsymbol{N}(\overline{\boldsymbol{\beta}}_p, \overline{\boldsymbol{V}}_p^{\beta}),$$

with

$$\overline{\boldsymbol{V}}_p^{\beta} = \left(\mathcal{X}'(\boldsymbol{V}_p^y)^{-1}\mathcal{X} + (\boldsymbol{V}_p^{\beta})^{-1}\right)^{-1},$$

and

$$\overline{\boldsymbol{\beta}}_p = \overline{\boldsymbol{V}}_p^{\beta} \times \left(\mathcal{X}'(\boldsymbol{V}_p^y)^{-1}\tilde{\boldsymbol{y}}\right),$$

where $\tilde{\boldsymbol{y}} = (\boldsymbol{y} - \theta \boldsymbol{z}_p)$.

2. The posterior inference for λ_p^2

Let $\beta_{i,p}$ denotes the ith component of $\boldsymbol{\beta}_p^{\Delta}$ for $i = 1, \ldots, TN$. The prior distribution for $\boldsymbol{\beta}_p^{\Delta}$ is actually equivalent to:

$$\beta_{i,p}|\lambda_p^2, \psi_{i,p}^2 \;\sim\; N(0, \lambda_p^2 \psi_{i,p}^2).$$

Therefore, the conditional posterior distribution for λ_p^2 has the form:

$$\begin{aligned} p(\lambda_p^2|\bullet) &\propto (\lambda_p^2)^{-\frac{TN}{2}} \exp\left\{-\left(\frac{1}{2}\sum_{i=1}^{TN}\frac{\beta_{i,p}^2}{\psi_{i,p}^2}\right)\frac{1}{\lambda_p^2}\right\} \times (\lambda_p^2)^{-\frac{1}{2}-1}\exp\left\{-\frac{1}{\xi_p}\frac{1}{\lambda_p^2}\right\} \\ &\propto (\lambda_p^2)^{-\frac{TN}{2}-\frac{1}{2}-1}\exp\left\{-\left(\frac{1}{2}\sum_{i=1}^{TN}\frac{\beta_{i,p}^2}{\psi_{i,p}^2} + \frac{1}{\xi_p}\right)\frac{1}{\lambda_p^2}\right\}. \end{aligned}$$

Therefore, the conditional posterior distribution for λ_p^2 is given by:

$$\lambda_p^2|\bullet \;\sim\; \mathcal{IG}\left(\frac{1+TN}{2}, \frac{1}{\xi_p} + \sum_{i=1}^{TN} \frac{\beta_{i,p}^2}{2\psi_{i,p}^2}\right). \tag{3.5}$$

3. The posterior inference for ξ_p

Since the ξ_p is the auxiliary variable for λ_p^2, it is easy to check that:

$$\begin{aligned} p(\xi_p|\bullet) &\propto (\xi_p)^{-\frac{1}{2}} \exp\left\{-\frac{1}{\lambda_p^2}\frac{1}{\xi_p}\right\} \times (\xi_p)^{-\frac{1}{2}-1} \exp\left\{-\frac{1}{\xi_p}\right\} \\ &\propto (\xi_p)^{-1-1} \exp\left\{-\left(1+\frac{1}{\lambda_p^2}\right)\frac{1}{\xi_p}\right\}. \end{aligned}$$

Therefore:

$$\xi_p|\bullet \;\sim\; \mathcal{IG}\left(1, 1+\frac{1}{\lambda_p^2}\right). \tag{3.6}$$

4. The posterior inference for $\psi_{i,p}^2$

The derivation of conditional posterior distribution for $\psi_{i,p}^2$ ($i = 1, \ldots, TN$) is similar with λ_p^2:

$$\begin{aligned} p(\psi_{i,p}^2|\bullet) &\propto (\psi_{i,p}^2)^{-\frac{1}{2}} \exp\left\{-\left(\frac{1}{2}\frac{\beta_{i,p}^2}{\lambda_p^2}\right)\frac{1}{\psi_{i,p}^2}\right\} \times (\psi_{i,p}^2)^{-\frac{1}{2}-1} \exp\left\{-\frac{1}{\zeta_{i,p}}\frac{1}{\psi_{i,p}^2}\right\} \\ &\propto (\psi_{i,p}^2)^{-\frac{1}{2}-\frac{1}{2}-1} \exp\left\{-\left(\frac{1}{2}\frac{\beta_{i,p}^2}{\lambda_p^2} + \frac{1}{\zeta_{i,p}}\right)\frac{1}{\psi_{i,p}^2}\right\}, \end{aligned}$$

so that:

$$\psi_{i,p}^2|\bullet \;\sim\; \mathcal{IG}\left(1, \frac{1}{\zeta_{i,p}} + \frac{\beta_{i,p}^2}{2\lambda_p^2}\right). \tag{3.7}$$

5. The posterior inference for $\zeta_{i,p}$

Similarly, it is easy to check that:

$$
\begin{aligned}
p(\zeta_{i,p}|\bullet) &\propto (\zeta_{i,p})^{-\frac{1}{2}} \exp\left\{-\frac{1}{\psi_{i,p}^2}\frac{1}{\zeta_{i,p}}\right\} \times (\zeta_{i,p})^{-\frac{1}{2}-1} \exp\left\{-\frac{1}{\zeta_{i,p}}\right\} \\
&\propto (\zeta_{i,p})^{-1-1} \exp\left\{-\left(1+\frac{1}{\psi_{i,p}^2}\right)\frac{1}{\zeta_{i,p}}\right\},
\end{aligned}
$$

then

$$
\zeta_{i,p}|\bullet \;\sim\; \mathcal{IG}\left(1, 1+\frac{1}{\psi_{i,p}^2}\right). \tag{3.8}
$$

6. The posterior inference for $z_{t,p}$

The derivation of the conditional posterior distribution for $z_{t,p}$ is the same as in section 2.2 and has similar form with Equ.2.13. Specifically, we have:

$$
z_{t,p}|\bullet \;\sim\; GIG\left(\frac{1}{2}, \frac{|y_t - \mathcal{X}_t\boldsymbol{\beta}_p^{\Delta}|}{\sigma_p\tau}, \frac{\sqrt{2\tau^2+\theta^2}}{\sigma_p\tau}\right),
$$

which is equivalent to:

$$
z_{t,p}^{-1}|\bullet \;\sim\; IG\left(\frac{\sqrt{2\tau^2+\theta^2}}{|y_t - \mathcal{X}_t\boldsymbol{\beta}_p^{\Delta}|}, \frac{2\tau^2+\theta^2}{\sigma_p^2\tau^2}\right), \tag{3.9}
$$

where $\mathcal{X}_t$ denotes the tth row of $\mathcal{X}$.

6. The posterior inference for σ_p^2

Given the hyperparameters $\underline{\rho}_1$ and $\underline{\rho}_2$, the conditional posterior distribution for σ_p^2 can be written as:

$$
\begin{aligned}
p(\sigma_p^2|\bullet) \propto\; &(\sigma_p^2)^{-\frac{T}{2}} \exp\left\{-\left(\sum_{t=1}^{T}\frac{(y_t - \mathcal{X}_t\boldsymbol{\beta}_p^{\Delta} - \theta z_t)}{2z_t\tau^2}\right)\frac{1}{\sigma_p^2}\right\}(\sigma_p^2)^{-T}\exp\left\{-\left(\sum_{t=1}^{T} z_t\right)\frac{1}{\sigma_p^2}\right\} \\
&\times (\sigma_p^2)^{-\underline{\rho}_1-1}\exp\left\{-\underline{\rho}_2\frac{1}{\sigma_p^2}\right\} \\
\propto\; &(\sigma_p^2)^{-\underline{\rho}_1-\frac{3T}{2}-1}\exp\left\{-\left(\underline{\rho}_2 + \sum_{t=1}^{T}\frac{(y_t - \mathcal{X}_t\boldsymbol{\beta}_p^{\Delta} - \theta z_t)}{2z_t\tau^2} + \sum_{t=1}^{T} z_t\right)\frac{1}{\sigma_p^2}\right\}.
\end{aligned}
$$

Therefore, we have:

$$\sigma_p^2|\bullet \sim \mathcal{IG}\left(\underline{\rho}_1 + \frac{3T}{2}, \underline{\rho}_2 + \sum_{t=1}^{T} \frac{(y_t - \mathcal{X}_t\boldsymbol{\beta}_p^{\Delta} - \theta z_t)}{2 z_t \tau^2} + \sum_{t=1}^{T} z_t\right). \tag{3.10}$$

3.4 Efficient high-dimensional sampling

Note that $\boldsymbol{\beta}_p^{\Delta}$ is a $TN \times 1$ high-dimensional vector, directly sample from a high-dimensional conditional posterior distribution is very time-consuming. Here we use the sampling algorithm provided by Bhattacharya *et al.* (2016). In our case, the sampling algorithm can be concluded by:

Step 1 Sample $\boldsymbol{\eta} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{V}_p^{\beta})$ and $\delta \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{I}_T)$

Step 2 Set $\boldsymbol{v} = \tilde{\mathcal{X}}\boldsymbol{\eta} + \boldsymbol{\delta}$

Step 3 Set $\boldsymbol{w} = (\tilde{\mathcal{X}}\boldsymbol{V}_p^{\beta}\tilde{\mathcal{X}}' + \boldsymbol{I}_T)^{-1}[\boldsymbol{y}^* - \boldsymbol{v}]$

Step 4 Set $\boldsymbol{\beta}_p^{\Delta} = \boldsymbol{\eta} + \boldsymbol{V}_p^{\beta}\tilde{\mathcal{X}}'\boldsymbol{w}$,

where $\tilde{\mathcal{X}} = (\boldsymbol{V}_p^{y})^{-\frac{1}{2}}\mathcal{X}$, $(\boldsymbol{V}_p^{y})^{-\frac{1}{2}}$ is a $T \times T$ diagonal matrix with tth diagonal element be $(\sigma_p \tau \sqrt{z_{t,p}})^{-1}$, and tth element of $\boldsymbol{y}^{\star} = (y_t - \theta z_{t,p})/(\sigma_p \tau \sqrt{z_{t,p}})$.

3.5 The MCMC algorithm

Given the conditional posterior distributions for $\boldsymbol{\beta}_p^{\Delta}, \lambda_p^2, \xi_p, \psi_{i,p}^2, \zeta_{i,p}, z_{t,p}, \sigma_p^2$ are all in closed-form. We implement the Gibbs sampler by recursively sampling parameters from:

1. Sample $\boldsymbol{\beta}_p^{\Delta}$ conditional on other parameters and the data:

$$\boldsymbol{\beta}_p^{\Delta}|\bullet \sim \boldsymbol{N}(\overline{\boldsymbol{\beta}}_p, \overline{\boldsymbol{V}}_p^{\beta}),$$

where $\overline{\boldsymbol{V}}_p^{\beta} = \left(\mathcal{X}'(\boldsymbol{V}_p^{y})^{-1}\mathcal{X} + (\boldsymbol{V}_p^{\beta})^{-1}\right)^{-1}$, $\overline{\boldsymbol{\beta}}_p = \overline{\boldsymbol{V}}_p^{\beta} \times \left(\mathcal{X}'(\boldsymbol{V}_p^{y})^{-1}\tilde{\boldsymbol{y}}\right)$ and $\boldsymbol{V}_p^{\beta}$ be a diagonal matrix with ith diagonal element be $\lambda_p^2\psi_{i,p}^2$. The high-dimensional sampling problem is addressed in section 3.4.

2. Sample λ_p^2, conditional on other parameters and the data:

$$\lambda_p^2|\bullet \sim \mathcal{IG}\left(\frac{1+TN}{2}, \frac{1}{\xi_p} + \sum_{i=1}^{TN} \frac{\beta_{i,p}^2}{2\psi_{i,p}^2}\right),$$

where $\beta_{i,p}$ is the ith element of $\boldsymbol{\beta}_p^{\Delta}$.

3. Sample ξ_p, conditional on other parameters and the data:

$$\xi_p|\bullet \sim \mathcal{IG}\left(1, 1 + \frac{1}{\lambda_p^2}\right).$$

4. Sample $\psi_{i,p}^2$ for $i = 1, \ldots, TN$, conditional on other parameters and the data:

$$\psi_{i,p}^2|\bullet \sim \mathcal{IG}\left(1, \frac{1}{\zeta_{i,p}} + \frac{\beta_{i,p}^2}{2\lambda_p^2}\right).$$

5. Sample $\zeta_{i,p}$ conditional on other parameters and the data:

$$\zeta_{i,p}|\bullet \sim \mathcal{IG}\left(1, 1 + \frac{1}{\psi_{i,p}^2}\right).$$

6. Sample $z_{t,p}$ for $t = 1, \ldots, T$, conditional on other parameters and the data:

$$z_{t,p}^{-1}|\bullet \sim IG\left(\frac{\sqrt{2\tau^2+\theta^2}}{|y_t - \mathcal{X}_t\boldsymbol{\beta}_p^{\Delta}|}, \frac{2\tau^2+\theta^2}{\sigma_p^2\tau^2}\right),$$

where $\mathcal{X}_t$ denotes the tth row of $\mathcal{X}$.

7. Sample σ_p^2, conditional on other parameters and the data:

$$\sigma_p^2|\bullet \sim \mathcal{IG}\left(\underline{\rho}_1 + \frac{3T}{2},\ \underline{\rho}_2 + \sum_{t=1}^{T} \frac{(y_t - \mathcal{X}_t\boldsymbol{\beta}_p^{\Delta} - \theta z_t)^2}{2z_t\tau^2} + \sum_{t=1}^{T} z_t\right).$$

4 Empirical study

4.1 The data

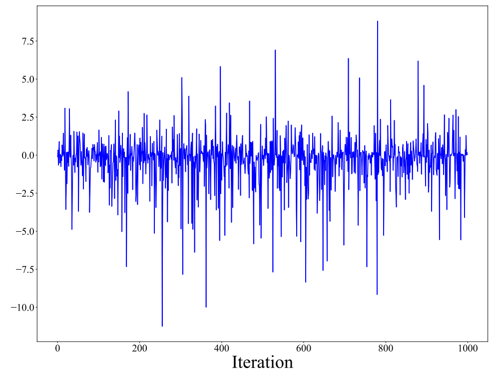
Table 1: Data Description

Mnemonic	Description	Source	Range
CPI	Consumer Price Index, Quarterly Vintages	Philly	1947Q1-2015Q3
IPM	Industrial Production Index, Manufacturing	Philly	1947Q1-2015Q3
HSTARTS	Housing Starts	Philly	1947Q1-2015Q3
CUM	Capacity Utilization Rate, Manufacturing	Philly	1948Q1-2015Q3
M1	M1 Money Stock	Philly	1947Q1-2015Q3
RCOND	Real Personal Consumption Expenditures, Durables	Philly	1947Q1-2015Q3
RCONS	Real Personal Consumption Expenditures, Services	Philly	1947Q1-2015Q3
RG	Real Government Consumption & Gross Investment, Total	Philly	1947Q1-2015Q3
RINVBF	Real Gross Private Domestic Investment, Nonresidential	Philly	1947Q1-2015Q3
ROUTPUT	Real GNP/GDP	Philly	1947Q1-2015Q3
RUC	Unemployment Rate	Philly	1948Q1-2015Q3
ULC	Unit Labor Costs	Philly	1947Q1-2015Q3
WSD	Wage and Salary Disbursements	Philly	1947Q1-2015Q3
DYS	Default yield spread (Moody's BAA - AAA)	St Louis	1947Q1-2015Q3
NAPM	Purchasing Manager's Index	St Louis	1947Q1-2015Q3
NAPMII	Inventories Index	St Louis	1947Q1-2015Q3
NAPMNOI	New Orders Index	St Louis	1947Q1-2015Q3

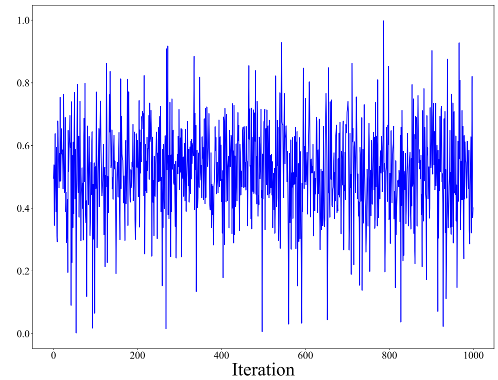
The data I use in this paper is quarterly macro data from 1948:Q1 to 2015: Q4. There are 274 observations. The description of data can be found in table 1.

4.2 BQR-SSVS

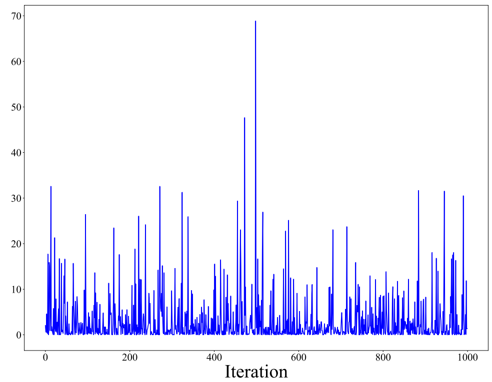
I first estimate the BQR-SSVS developed in section 2. The dependent variable is CPI, and the covariates include the AR(2) of CPI, and all the other 16 macro indicators. I take the intercept into consideration, so there are 19 variables in total in the covariates.



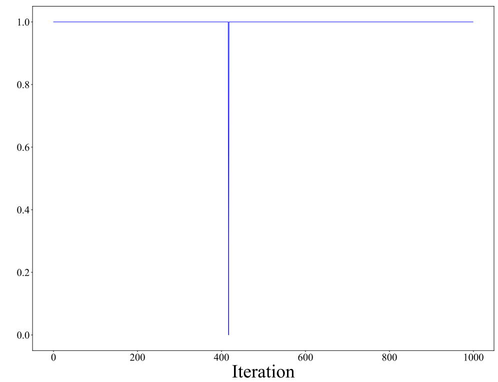
(a) Trace Plot for $\beta_{0,0.05}$



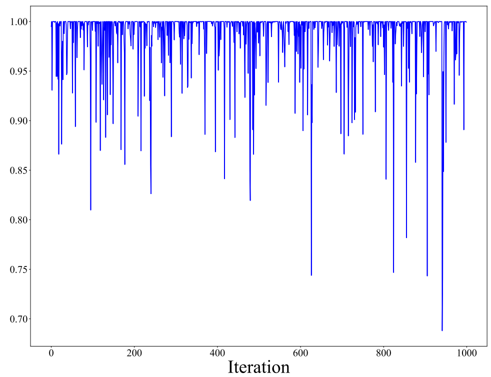
(b) Trace Plot for $\beta_{1,0.05}$



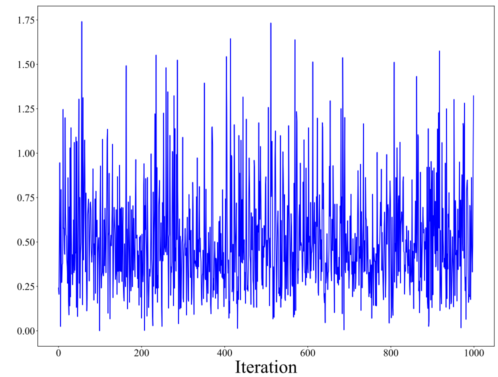
(c) Trace Plot for $\delta_{0,0.05}^{-2}$



(d) Trace Plot for $\gamma_{0,0.05}$



(e) Trace Plot for $\pi_{0,0.05}$



(f) Trace Plot for $z_{0,0.05}$

Figure 1: Trace plots for the MCMC estimation of BQR-SSVS.

I consider $q = 0.05, 0.15, \ldots, 0.95$, so there are 10 different quantiles. I set $\underline{a}_1 = \underline{a}_2 = 0.1$, so that the prior mean for δ^{-2} equals to 1, and the prior variance for δ^{-2} equals to 10. Moreover, I set $\underline{b}_1 = \underline{b}_2 = 0.1$, so that the prior mean for π_0 equals to 0.5, and the prior variance is approximately 0.25. I let $\underline{c} = 10^{-}3$, a very small number. Following Korobilis (2017), 90000 Monte Carlo iterations have been used, 50000 iterations are discarded for convergence, and from the remaining 40000 draws only every 40th draw is retained.

Figures 1(a) to 1(f) show the trace plots (partly) of MCMC estimation of BQR-SSVS. $\beta_{0,0.05}$ denotes the trace plot for coefficient of the intercept at 0.05 quantile, and $\beta_{1,0.05}$ denotes the trace plot for coefficient of the AR(1) term of CPI at 0.05 quantile. I highlight three findings. First, in general, the sampled sequences have no obvious auto-correlations, indicating that the convergence of Gibbs sampler. Second, the sampling results of the coefficient of intercept fluctuates more than the first-order lag term, indicating that the intercept may need to be modeled independently. Third, the samples of $\gamma_{0,0.05}$ are always equal to 1, implying that the SSVS does not want β_0 to be zero.

I then compare the model performance of the BQR-SSVS with frequentist approach, non-parametric estimation of the quanile regression[1]. The results is shown in figure 2. For each sub-plot, I plot the estimated coefficient across 10 different quantiles from 0.05 to 0.95. The blue point denotes the posterior mean and the red line denotes the 95% confidence interval estimated by BQR-SSVS. Similarly, the orange point and green line denote the mean and 95% confidence interval estimated by frequentist approach. It can be found from the results that except for the intercept, the estimation results of BQR-SSVS and the frequentist's approach are consistent in trend and magnitude. Moreover, the coefficients are also significant different across different quantiles, which also reflects the robustness of quantile regression compared to mean regression.

[1] I use the Python package QuantReg in statsmodels.

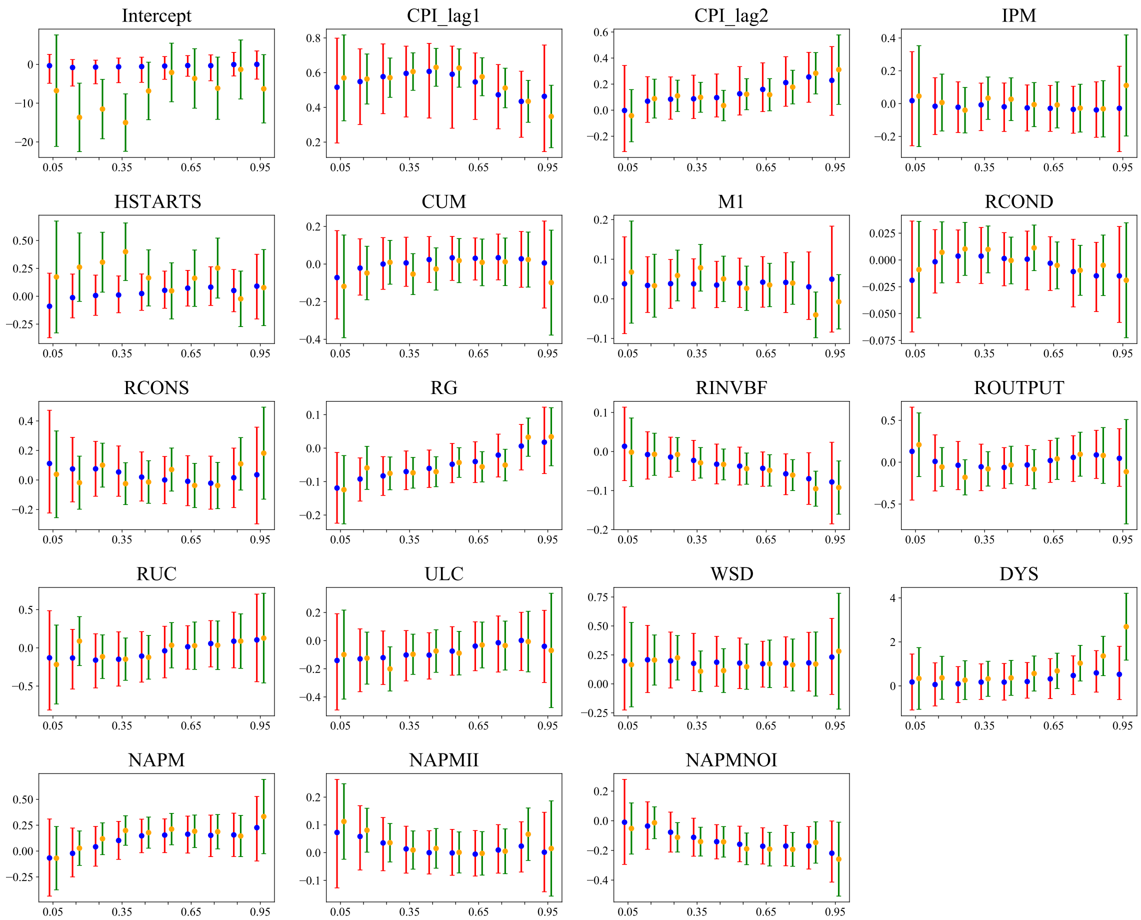


Figure 2: BQR-SSVS verses frequentist approach.

4.3 TVP-BQR

Now let's move into the TVP-BQR model. It can be quickly recognized that if I replace the $\boldsymbol{\mathcal{X}}_t$ to the covariates $\boldsymbol{x}_t$ in Equ.(3.4), the the estimated $\boldsymbol{\beta}^{\Delta}$ has the same interpretation of $\boldsymbol{\beta}$ in the previous section. Referring to Kozumi and Kobayashi (2011), I set $\rho_1 = 2.5$ and $\rho_2 = 0.5$. 70000 Monte Carlo iterations have been used, 50000 iterations are discarded for convergence.

The estimated coefficients are shown in figure 3. Based on the results in figure 2, I add the posterior mean (purple point) and 95% confidence interval (yellow line) estimated by the proposed BQR with horseshoe prior. I find that in general, the estimated mean and confidence level of the three methods are very similar. However, these methods are still inconsistent in the coefficient of the intercept. An interesting finding is the the BQR with horseshoe prior seems to have tighter confidence interval. This may be due to the fact that the horseshoe prior has the strongest prior shrinkage.

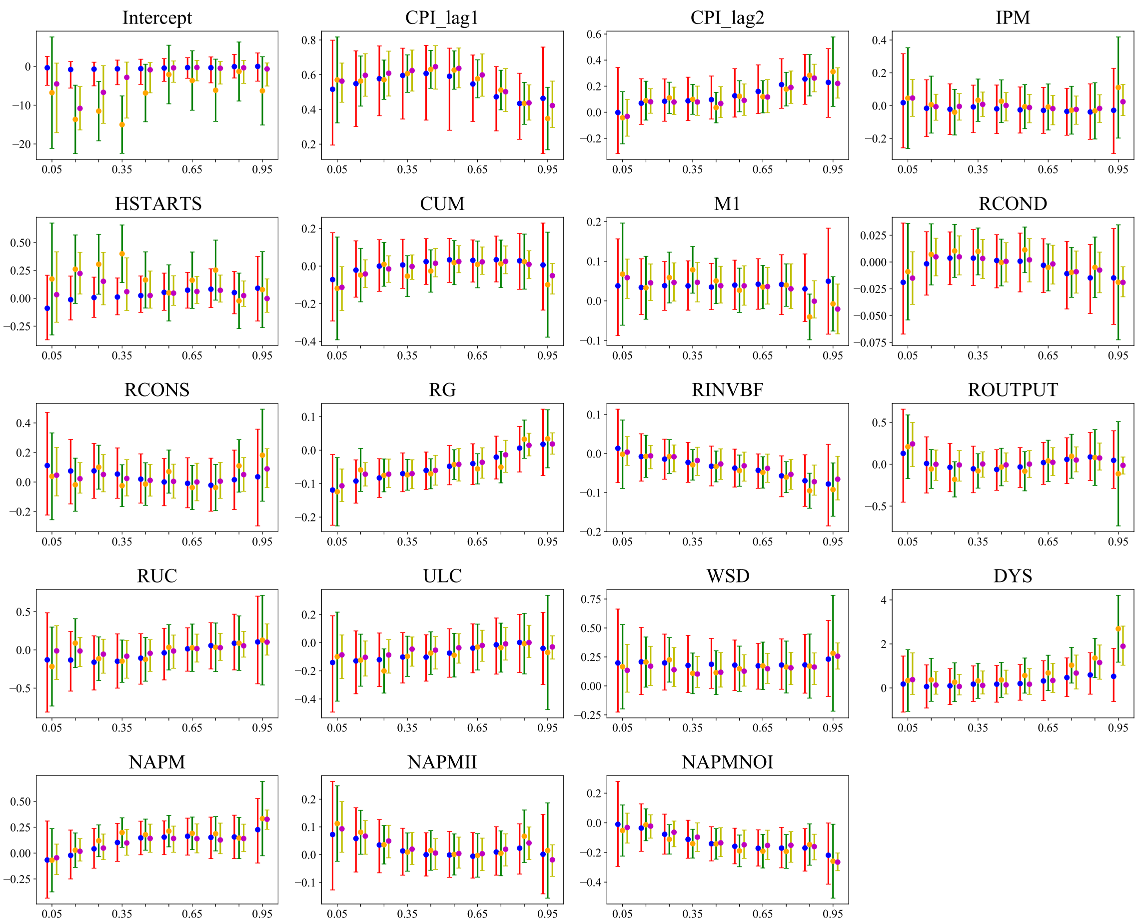


Figure 3: Performance of TVP-BQR.

Finally, I would like to access the time-varying coefficients of the AR(2) specification of CPI. Let y be the CPI, the model I consider is:

$$y_t = \beta_{0,t}(p) + \beta_{1,t}(p)y_{t-1} + \beta_{2,t}(p)y_{t-2} + \epsilon_t.$$

The results are plotted in figure 4. I highlight two findings. First, the estimated coefficients show strong time-varying effects, and there were significant structural changes in the 2008 global financial crisis. Second, for the lag terms of CPI, it can be found that the coefficients change more sharply in the 0.05 quantile compared to the median or 0.95 quantile, indicating the importance and necessity of measuring the tail performance.

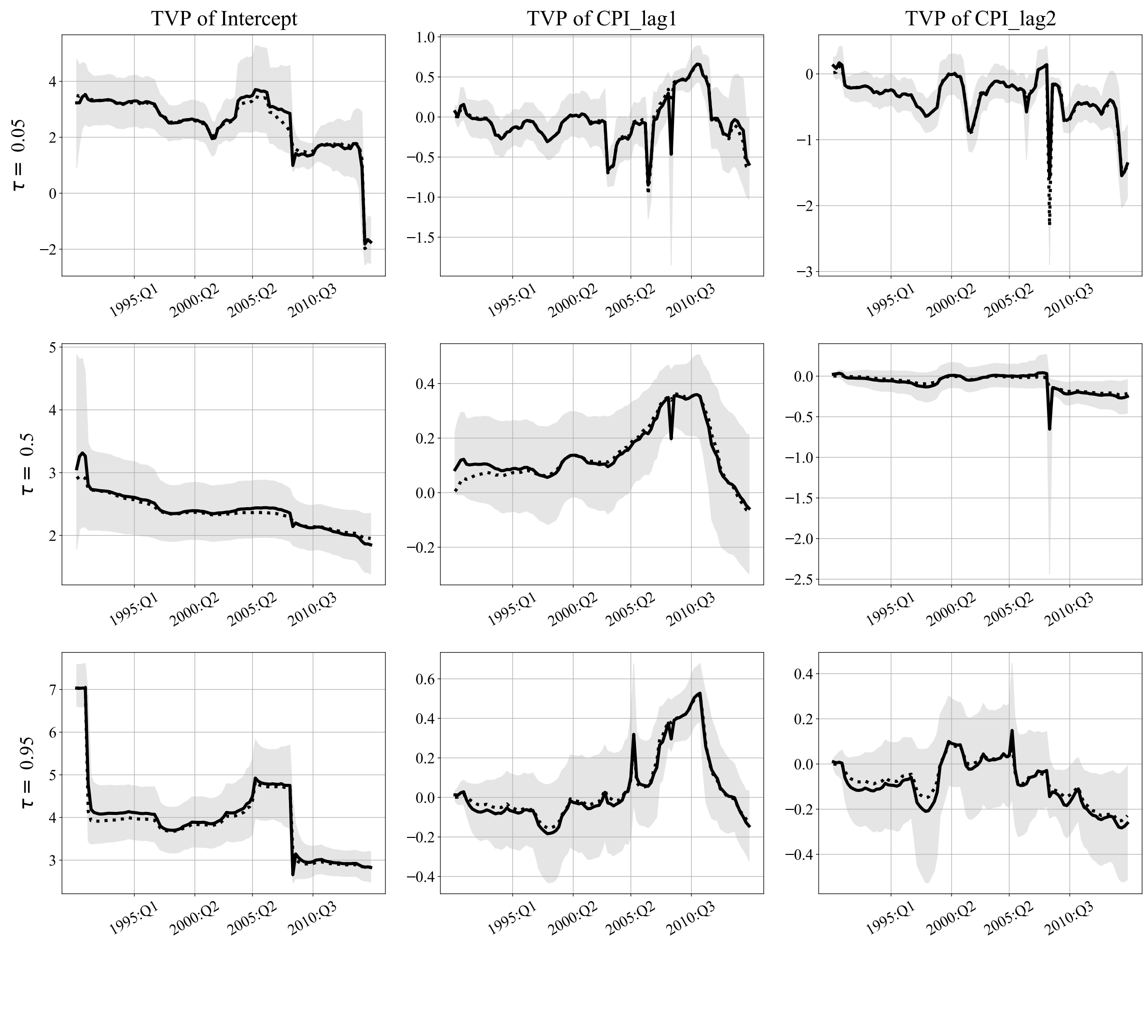


Figure 4: Estimated coefficients of TVP-BQR.

References

Benati L, and Surico P (2009), "VAR analysis and the Great Moderation," *American Economic Review* **99**(4), 1636–52.

Bernanke B (2004), "The great moderation," *Washington, DC* .

Bhattacharya A, Chakraborty A, and Mallick BK (2016), "Fast sampling with Gaussian scale mixture priors in high-dimensional regression," *Biometrika* asw042.

Carvalho CM, Polson NG, and Scott JG (2010), "The horseshoe estimator for sparse signals," *Biometrika* **97**(2), 465–480.

Durbin J, and Koopman SJ (2012), *Time series analysis by state space methods*, volume 38, OUP Oxford.

Hong Y, Wang X, and Wang S (2017), "Testing strict stationarity with applications to macroeconomic time series," *International Economic Review* **58**(4), 1227–1277.

Justiniano A, and Primiceri GE (2008), "The time-varying volatility of macroeconomic fluctuations," *American Economic Review* **98**(3), 604–41.

Khare K, and Hobert JP (2012), "Geometric ergodicity of the Gibbs sampler for Bayesian quantile regression," *Journal of Multivariate Analysis* **112**, 108–116.

Kim CJ, and Nelson CR (1999), "Has the US economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle," *Review of Economics and Statistics* **81**(4), 608–616.

Koenker R, and Bassett Jr G (1978), "Regression quantiles," *Econometrica: journal of the Econometric Society* 33–50.

Korobilis D (2017), "Quantile regression forecasts of inflation under model uncertainty," *International Journal of Forecasting* **33**(1), 11–20.

——— (2021), "High-dimensional macroeconomic forecasting using message passing algorithms," *Journal of Business & Economic Statistics* **39**(2), 493–504.

Korobilis D, Landau B, Musso A, and Phella A (2021), "The time-varying evolution of inflation risks," .

Kozumi H, and Kobayashi G (2011), "Gibbs sampling methods for Bayesian quantile regression," *Journal of statistical computation and simulation* **81**(11), 1565–1578.

Makalic E, and Schmidt DF (2015), "A simple sampler for the horseshoe estimator," *IEEE Signal Processing Letters* **23**(1), 179–182.

Rossi B (2013), "Exchange rate predictability," *Journal of economic literature* **51**(4), 1063–1119.

Stock JH, and Watson MW (1996), "Evidence on structural instability in macroeconomic time series relations," *Journal of Business & Economic Statistics* **14**(1), 11–30.

Welch I, and Goyal A (2008), "A comprehensive look at the empirical performance of equity premium prediction," *The Review of Financial Studies* **21**(4), 1455–1508.

Yu K, and Moyeed RA (2001), "Bayesian quantile regression," *Statistics & Probability Letters* **54**(4), 437–447.