Open access • Journal Article • DOI:10.1111/STAN.12091

Markov switching quantile autoregression — Source link

Xiaochun Liu

Institutions: University of Alabama

Published on: 01 Nov 2016 - Statistica Neerlandica (John Wiley & Sons, Ltd)

Topics: Quantile, Markov chain, Bayesian inference, Markov process and Asymmetric Laplace distribution

Markov-Switching Quantile Autoregression

Liu, Xiaochun

Department of Economics, Emory University

7 October 2013

Markov-Switching Quantile Autoregression

Xiaochun Liu*

May 7, 2014

Abstract

This paper considers the location-scale quantile autoregression in which the location and scale parameters are subject to regime shifts. The regime changes are determined by the outcome of a latent, discrete-state Markov process. The new method provides direct inference and estimate for different parts of a nonstationary time series distribution. Bayesian inference for switching regimes within a quantile, via a three-parameter asymmetric-Laplace distribution, is adapted and designed for parameter estimation. The simulation study shows reasonable accuracy and precision in model estimation. From a distribution point of view, rather than from a mean point of view, the potential of this new approach is illustrated in the empirical applications to reveal the countercyclical risk pattern of stock markets and the asymmetric persistence of real GDP growth rates and real trade-weighted exchange rates.

Keywords: Asymmetric-Laplace Distribution, Metropolis-Hastings, Block-at-a-Time, Asymmetric Dynamics, Transition Probability

JEL: C22, C58, C51, C11, G23

*Department of Economics, Emory University, Atlanta GA 30322. Email: xiaochun.liu@emory.edu

1 Introduction

Koenker and Xiao (2006) study quantile autoregression models in which the autoregressive coefficients may take distinct values over different quantiles of the innovation process. Their models can capture systematic influences of conditioning variables on the location, scale and shape of the conditional distribution. Let $\{U_t\}$ be a sequence of i.i.d. standard uniform random variables. Consider the mth-order autoregressive process

$$y_t = \theta_0(U_t) + \theta_1(U_t)\, y_{t-1} + ... + \theta_m(U_t) y_{t-m} \tag{1.1}$$

where y_t is the time series observation at time t, and θ's are unknown functions $[0,1] \to \mathbb{R}$ to be estimated. Provided that the right side of (1.1) is monotone increasing in U_t, it flows that the τth conditional quantile function of y_t can be obtained as

$$Q_{y_t}\left(\tau | \boldsymbol{y}_{t-1}\right) = \theta_0(\tau) + \theta_1(\tau) y_{t-1} + ... + \theta_m(\tau) y_{t-m} \tag{1.2}$$

where $\boldsymbol{y}_{t-1} = (y_{t-1}, ..., y_{t-m})'$. The transition from (1.1) to (1.2) is an immediate consequence of equivariance to monotone transformations.[1] In (1.2), the quantile autoregressive coefficients may be τ-dependent and thus can vary over the quantiles. The conditioning variables not only shift the location of the distribution of y_t , but also may alter the scale and shape of the conditional distribution. Koenker and Xiao (2006) also show that quantile autoregressive models exhibit a form of asymmetric persistence and temporarily explosive behavior.

However, the linear quantile autoregressive models cannot accommodate many stylized facts such as structural breaks and nonlinearities in macroeconomic and financial time series. The aim of this article is to extend the quantile autoregression of Koenker and Xiao (2006) by modeling nonstationary quantile dynamics. Particularly, I consider the location-scale quantile autoregression in which the location and scale parameters are subject to regime shifts within a quantile. Switching quantile regimes is determined by the outcome of an

[1] See the theorem of equivariance to monotone transformations in Koenker (2005), page 39.

unobserved state indicator variable that follows a Markov process with unknown transition probabilities. The proposed Markov-Switching Quantile Autoregression (MSQAR) nests the quantile autoregression of Koenker and Xiao (2006) as a special case when conditional distributions are stationary.

MSQAR is a convenient approach built on the vast literature of Markov-switching time series models.[2] Nonetheless, simply combining quantile autoregressive models with Markov-switching techniques is econometrically infeasible. The challenge is that the objective function of quantile autoregression is a non-likelihood based function generally estimated by nonlinear least square. The non-likelihood based function does not allow make inference on the latent state variable for switching regimes. To solve this problem, I assume that quantile error terms follow a three-parameter asymmetric-Laplace distribution (Yu and Zhang (2005)). This paper shows that maximizing this distribution is mathematically equivalent to minimizing quantile objective functions. Importantly, it also satisfies the restrictive conditions of quantile regression. With this distribution, the inference for switching quantile regimes can be made through the standard Hamilton filter approach (Hamilton (1994)).

This paper adopts Bayesian approach for model estimation. As discussed in Yu and Moyeed (2001), the use of an asymmetric Laplace distribution for error terms provides a natural way to deal with some serious computational challenges through Bayesian quantile regression. Also see Chernozhukov and Hong (2003). In the terminology of Chib and Greenberg (1995), this paper adopts a "block-at-a-time" Metropolis-Hastings sampling to reduce computational cost. This algorithm groups highly correlated parameters as one block to be simultaneously updated at each Metropolis-Hasting step conditional on the remaining blocks, see e.g., Tierney (1994), Ausin and Lopes (2010), Geweke and Tanizaki (2001), among others. To further speed up convergence and to achieve desirable mixing properties in MCMC chains, I employ the adaptive scheme of Gerlach et al. (2011) and Chen et al. (2012), which combines a random walk and an independent kernel Metropolis-

[2]See e.g., Sims and Zha, 2006, Gray (1996), Cheung and Erlandsson (2005), Hamilton and Susmel (1994), Kim et al. (2008), among many others. Guidolin (2012) provides a recent review for the applications of Markov-switching models in empirical finance.

Hastings algorithm, each based on a mixture of multivariate normal distributions.

This paper examines the new approach in a simulation study to show its accuracy and precision in model estimation. The empirical application to S&P 500 returns illustrates the usefulness of this new approach in risk management, i.e., for stress-testing financial institutions from the perspective of central banks. In this paper, asymmetric dynamics have also been found for quarterly real GDP growth rates but not for quarterly real trade-weighted U.S. dollars. In addition, the asymmetric dynamics appear to be differernt across economic regimes. Notably, modeling the regime persistence in lower tails of real GDP growth rates improves the predictabilities of swtiching economic states and turning points.

The rest of this paper is structured as follows. Section 2 introduces the connection of asymmetric-Laplace distributions to the solution of quantile regressions. Section 3 defines Markov-Switching quantile autoregression. Section 4 describes the Bayesian methods in this paper for model estimation. Section 5 presents model simulations and results. Section 6 reports the results of empirical applications to stock markets, real GDP growth rates and real trade-weighted exchange rates. Section 7 concludes this paper.

2 Asymmetric Laplace Distribution Connection

The QAR(m) model of (1.2) can be reformulated in a more conventional regression form as

$$y_t = \theta_0(\tau) + \sum_{l=1}^{m} \theta_l(\tau) y_{t-l} + \varepsilon_t(\tau) \tag{2.1}$$

where $\varepsilon_t(\tau)$ is quantile error terms which follow an asymmetric-Laplace (AL) distribution, denoted by $AL(0, \varsigma, \tau)$, with the density function as

$$f_{\varepsilon_t}(\varepsilon; 0, \varsigma, \tau) = \frac{\tau(1-\tau)}{\varsigma} exp\left\{ -\frac{\varepsilon\left(\tau - I(\varepsilon \leq 0)\right)}{\varsigma} \right\} \tag{2.2}$$

where $I(\cdot)$ is an indicator function. τ determines the skewness of the distribution, $\varsigma > 0$ is a scale parameter. $AL(0, \varsigma, \tau)$ with the location parameter being zero provides that the τth

quantile of the distribution is zero as $Pr\left(\varepsilon_t \leq 0\right) = \tau$, which satisfies the quantile regression condition $\int_{-\infty}^{0} f_{\varepsilon}(s)ds = \tau$. The asymmetric-Laplace distribution with the density function of (2.2) has the mean and variance, $E(\varepsilon_t) = \varsigma(1-2\tau)/[(1-\tau)\tau]$ and $Var(\varepsilon_t) = \varsigma^2(1-2\tau+2\tau^2)/[(1-\tau)^2\tau^2]$, respectively. See Yu and Zhang (2005) for details. With the assumption of i.i.d. $\varepsilon_t(\tau)$, the sample likelihood function is given by

$$\begin{aligned} L(\boldsymbol{\theta}, \tau) &= [\tau(1-\tau)/\varsigma]^T \\ & \quad exp\left\{ -\sum_{t=1}^{T} \frac{y_t - Q_{y_t}\left(\tau|\boldsymbol{y}_{t-1}\right)}{\varsigma} \left[\tau - I\left(y_t \leq Q_{y_t}\left(\tau|\boldsymbol{y}_{t-1}\right)\right)\right] \right\} \end{aligned} \tag{2.3}$$

In the literature the error density is often left unspecified, see e.g., Koenker and Bassett (1978), Koenker (2005), and Koenker and Xiao (2006), etc. Quantile autoregression is the solution to the following minimization problem

$$\boldsymbol{\theta}(\tau) = arg\ \min_{\boldsymbol{\theta}} E\left(\rho_\tau\left(y_t - Q_{y_t}\left(\tau|\boldsymbol{y}_{t-1};\boldsymbol{\theta}\right)\right)\right) \tag{2.4}$$

where $\theta\left(\tau\right) = \left(\theta_0(\tau), ..., \theta_m\left(\tau\right)\right)$ is the parameter space to be estimated. The quantile criterion (check or loss) function $\rho_\tau(\cdot)$ is defined as $\rho_\tau(\varepsilon) = \varepsilon\left(\tau - I(\varepsilon < 0)\right)$ in Koenker and Bassett (1978). Solving the sample analog gives the estimator of θ

$$\hat{\boldsymbol{\theta}}(\tau) = arg\ \min_{\boldsymbol{\theta}} \sum_{t=1}^{T} \rho_\tau\left(y_t - Q_{y_t}\left(\tau|\boldsymbol{y}_{t-1};\boldsymbol{\theta}\right)\right) \tag{2.5}$$

Recently, Yu and Moyeed (2001), Yu and Zhang (2005) and Gerlach et al. (2011), among others, have illustrated the link between the quantile estimation problem and asymmetric-Laplace distribution. Since the quantile loss function is contained in the exponent of the asymmetric-Laplace likelihood, maximizing the sample likelihood of (2.3) is mathematically equivalent to minimizing the quantile loss function of (2.5). It is important to emphasize that, in practice, the parameter τ is chosen by researchers as quantile levels of interest during parameter estimation and only a single quantile of the distribution of y_t is estimated. More importantly, the asymmetric Laplace distribution transforms the non-likelihood based

quantile regression of (2.5) to a likelihood based approach, so that the inference for the probability of switching regimes is possibly made through Hamilton filter.

3 Markov-Switching Quantile Autoregression

For the τth conditional quantile of y_t, let $\{s_t\}$ be an ergodic homogeneous Markov chain on a finite set $S = \{1, ..., k\}$, with a transition matrix P defined by the following transition probabilities

$$\{p_{ij} = Pr\left(s_t = j | s_{t-1} = i\right)\}$$

and the unconditional probabilities

$$\{\pi_j = Pr\left(s_t = j\right)\}$$

for $i, j \in S$ and assume s_t follow a first-order Markov chain. The transition probabilities satisfy $\sum_{j \in S} p_{ij} = 1$ and $\sum_{j \in S} \pi_j = 1$. The stochastic process for s_t is strictly stationary if p_{ij} is less than unity and does not take on the value of 0 simultaneously.

Using transition probabilities above, this paper defines Markov-Switching quantile autoregressive models (MSQAR) as

$$\begin{aligned} y_t &= Q_{y_t}(\tau | \mathbf{y}_{t-1}; \boldsymbol{\theta}_{s_t}) + \varepsilon_t(\tau) \\ &= \theta_{s_t,0}(\tau) + \sum_{l=1}^{m} \theta_{s_t,l}(\tau) y_{t-l} + \varepsilon_t(\tau) \end{aligned} \tag{3.1}$$

Suppose that $\mathbf{y}_t$ can be observed directly but can only make an inference about the value of s_t based on the observations as of date t. This inference gives the filtering probability as

$$\begin{aligned} \xi_{j,t|t} &= Pr\left(s_t = j | \mathbf{y}_t; \boldsymbol{\Theta}\right) \\ &= \sum_{i \in S} Pr\left(s_t = j, s_{t-1} = i | \mathbf{y}_t; \boldsymbol{\Theta}\right) \end{aligned}$$

where $\sum_{j \in S} \xi_{j,t|t} = 1$ and $\boldsymbol{\Theta} = (P, \boldsymbol{\theta}_{s_t}(\tau))$ is a vector of the parameters with $s_t \in S$. The

formulation of filtering probabilities is obtained by Bayes theorem as

$$\xi_{j,t|t} = \frac{\sum_{i \in S} p_{ij} \xi_{i,t-1|t-1} \eta_{j,t}}{f\left(y_t | \mathbf{y}_{t-1}, \tau; \boldsymbol{\Theta}\right)} \tag{3.2}$$

where $\eta_{j,t}$ is conditional likelihood as

$$\begin{aligned} \eta_{j,t} &= f\left(y_t | s_t = j, \mathbf{y}_{t-1}, \tau; \boldsymbol{\theta}\right) \\ &= \frac{\tau(1-\tau)}{\varsigma} exp\left\{ -\frac{\left(y_t - Q_{y_t}\left(\tau | \mathbf{y}_{t-1}; \boldsymbol{\theta}_j\right)\right)}{\varsigma} \left[\tau - I\left(y_t < Q_{y_t}\left(\tau | \mathbf{y}_{t-1}; \boldsymbol{\theta}_j\right)\right)\right] \right\} \end{aligned} \tag{3.3}$$

and

$$f(y_t | \mathbf{y}_{t-1}, \tau; \boldsymbol{\Theta}) = \sum_{j \in S} \sum_{i \in S} p_{ij} \xi_{i,t-1|t-1} \eta_{j,t}$$

Thus, the relationship between the filtering and prediction probabilities is given by

$$\xi_{j,t+1|t} = Pr\left(s_{t+1} = j | \mathbf{y}_t; \boldsymbol{\Theta}\right) = \sum_{i \in S} p_{ij} \xi_{i,t|t} \tag{3.4}$$

The inference, similar to Hamilton's filter (Hamilton, 1994), is performed iteratively for $t = 1, ..., T$ with the initial values, $\xi_{j,0|0}$ for $j \in S$. The sample likelihood for the τth conditional quantile of y_t is then given by

$$L(\boldsymbol{\Theta}) = \prod_{t=1}^{T} f(y_t | \mathbf{y}_{t-1}, \tau; \boldsymbol{\Theta}) \tag{3.5}$$

The connection to the solution of quantile regression can also be viewed as follows. Based on quantile loss functions, $\boldsymbol{\Theta}$ is solved for the following minimization problem

$$\min_{\boldsymbol{\Theta}} E\left(\sum_{j \in S} \rho_\tau\left(y_t - Q_{y_t}\left(\tau | s_t = j,\ \mathbf{y}_{t-1}; \boldsymbol{\Theta}\right)\right) I\left(s_t = j\right) \right) \tag{3.6}$$

where $\mathbf{y}_t = \{y_t, y_{t-1}, ..., y_1, y_0\}$. Apply the law of iterated expectation to rewritten (3.6) as

$$\min_{\boldsymbol{\Theta}} \sum_{j \in S} E\left[\rho_\tau\left(y_t - Q_{y_t}\left(\tau | s_{t,\tau} = j,\ \mathbf{y}_{t-1}; \boldsymbol{\Theta}\right)\right) Pr\left(s_{t,\tau} = j | \mathbf{y}_t; \boldsymbol{\Theta}\right) \right] \tag{3.7}$$

Provided that τ is chosen by researchers of interest, maximizing the likelihood of (3.5) is mathematically equivalent to the minimization of (3.6), since the likelihood function can be alternatively rewritten as $L(\mathbf{\Theta}) = \prod_{t=1}^{T} \sum_{j \in S} f\left(y_t | s_t = j, \mathbf{y}_{t-1}, \tau; \mathbf{\Theta}\right) Pr\left(s_t = j | \mathbf{y}_t; \mathbf{\Theta}\right)$ with $Pr(s_t = j | \mathbf{y}_t; \mathbf{\Theta}) = \sum_{i \in S} p_{ij} \xi_{i,t-1|t-1}$. However, $Pr\left(s_t = j | \mathbf{y}_t; \mathbf{\Theta}\right)$ cannot be filtered by using the nonlinear least square estimation of (3.7); therefore, the likelihood function of the asymmetric Laplace distribution is used to infer transition probabilities.

To estimate smoothing transition probabilities $Pr\left(s_t = i | \mathbf{y}_T; \theta\right)$, this paper follows the approach of Kim (1994). Apply the Bayes theorem and the Markov property to yield

$$Pr\left(s_t = i | s_{t+1} = j,\ \mathbf{y}_T; \mathbf{\Theta}\right) = \frac{p_{ji} Pr\left(s_t = j | \mathbf{y}_t; \mathbf{\Theta}\right)}{Pr\left(s_{t+1} = i | \mathbf{y}_t; \mathbf{\Theta}\right)} \tag{3.8}$$

It is therefore the case that

$$Pr\left(s_t = j, s_{t+1} = i | \mathbf{y}_T; \mathbf{\Theta}\right) = Pr\left(s_{t+1} = i | \mathbf{y}_T; \mathbf{\Theta}\right) \frac{p_{ji} Pr\left(s_t = j | \mathbf{y}_t; \mathbf{\Theta}\right)}{Pr\left(s_{t+1} = i | \mathbf{y}_t; \mathbf{\Theta}\right)} \tag{3.9}$$

The smoothed inference for date t is the sum of (3.9) over $i \in S$

$$\begin{aligned} \xi_{j,t|T} &= Pr\left(s_t = j | \mathbf{y}_T; \mathbf{\Theta}\right) \\ &= \sum_{i \in S} Pr\left(s_{t+1} = i | \mathbf{y}_T; \mathbf{\Theta}\right) \frac{p_{ji} Pr\left(s_t = j | \mathbf{y}_t; \mathbf{\Theta}\right)}{Pr\left(s_{t+1} = i | \mathbf{y}_t; \mathbf{\Theta}\right)} \end{aligned} \tag{3.10}$$

The smoothed transition probabilities are thus obtained by iterating on (3.10) backward for $t = T-1, T-2, ..., 1$. This iteration is started with $\xi_{j,T|T}$ for $j \in S$ which is estimated from (3.2) for $t = T$. This algorithm is valid only when s_t follows a first-order Markov chain.

From the conditional density (3.3), it is straightforward to forecast the one-step-ahead τth quantile of y_{t+1} at time t conditional on knowing $s_{t+1,\tau}$,

$$Q_{y_{t+1}}\left(\tau | s_{t+1} = j, \mathbf{y}_t; \boldsymbol{\theta}_j\right) = \theta_{j,0}(\tau) + \sum_{l=0}^{m-1} \theta_{j,l+1}(\tau) y_{t-l} \tag{3.11}$$

Further, from (3.4), the forecast of $Q_{y_{t+1}}(\tau|\mathbf{y}_t;\boldsymbol{\theta}_j)$ is obtained as

$$Q_{y_{t+1}}(\tau|\mathbf{y}_t;\boldsymbol{\theta}_j) = \sum_{j=1}^{k} Q_{y_{t+1}}(\tau|s_{t+1}=j,\mathbf{y}_t;\boldsymbol{\theta}_j)\, Pr(s_{t+1}=j|\mathbf{y}_t,\boldsymbol{\Theta}) \tag{3.12}$$

which is to multiply the appropriate forecast of the quantile in the jth regime given by (3.11) with the probability that the process will be in that regime given by (3.4), and to sum those products for every regime together. Note that h-step-ahead forecasts for $h > 2$ require different approaches since it involves forecasts of y_{t+h-1} in (3.11) for $Q_{y_{t+h}}(\tau|s_{t+h-1,\tau}=j,\mathbf{y}_{t+h-1};\boldsymbol{\theta}_j)$, as shown in Cai (2010).

In MSQAR model estimation, similar to other Markov-Switching time series models, one must use some identification restrictions to avoid the label switching issue. See Bauwens et al. (2010) and Hamilton et al. (2007) for a discussion. In this paper, regimes are labeled by the restrictions on quantile intercepts, for example, $\theta_{1,0}(\tau) > ... > \theta_{k,0}(\tau)$. In addition, in empirical applications, the transition probabilities are allowed but not imposed dependent on τ. The intuition is that even though economic states are common across quantiles implying the same unconditional probabilities, no theories show that regime persistence should be the same across quantiles. To obtain some insights on this empirical question, regime persistence is allowed to be driven by data across quantiles.

4 Bayesian Inference

MSQAR models are non-linear and involve indicator functions, which introduce kinks and discontinuities into the sample likelihood function in (3.5). In addition, less observations fall in more extreme quantiles, which leads to the potential small sample issue. These issues make classical methods such as MLE very difficult for model estimation. In this paper, I instead prefer to use Bayesian MCMC methods to learn about the model parameters.

Given the sample realizations, y_t for $t = 1, ..., T$, the posterior distribution of $\boldsymbol{\Theta}$ takes the usual form: $p(\boldsymbol{\Theta}|\mathbf{y}_t) \propto L(\mathbf{y}_t|\boldsymbol{\Theta})\pi(\boldsymbol{\Theta})$, where $L(\mathbf{y}_t|\boldsymbol{\Theta})$ is the sample likelihood function and

$\pi(\boldsymbol{\Theta})$ is the prior distribution. Yu and Moyeed (2001) and Cai and Stander (2008) prove that the posterior distribution is proper under the improper prior for general quantile regression models. In this paper, the prior distribution is taken as uniform over Ξ, the admissible parameter space of $\boldsymbol{\Theta}$, i.e., satisfying the label switching restrictions. The prior for the scale parameter is $\pi(\varsigma) \propto \varsigma^{-1}$ also used in Gerlach et al. (2011).

Just like Vrontos et al. (2002) and Ausin and Lopes (2010), I also find that MCMC mixing can be improved and the computational cost reduced by using simultaneous updating of the highly correlated parameter groups at each Metropolis-Hastings (MH) step. In the terminology of Chib and Greenberg (1995), this approach is therefore based on a "block-at-a-time" MH sampler which is carried out by cycling repeatedly through draws of each parameter block conditional on the remaining parameter blocks. Let $\boldsymbol{\Theta} = (\boldsymbol{P}, \boldsymbol{\theta}_1(\tau), ..., \boldsymbol{\theta}_k(\tau))$ represent the blocks of the population parameters. $\boldsymbol{P} = (p_{ij})$ contains all transition probability parameters and $\boldsymbol{\theta}_{j,\tau}$ includes all parameters in the jth regime for $j = 1, ..., k$. Hence, the parameters in $\boldsymbol{\Theta}$ are grouped in $k+1$ blocks and the parameters of each block are simultaneously updated conditional on the remaining blocks.

This paper implements the MH sampler according to the adaptive scheme of Gerlach et al. (2011) and Chen et al. (2012) which combines the random walk MH (RW-MH) and the independent kernel MH (IK-MH) algorithms, each based on a mixture of multivariate normal distributions. The random walk part of this scheme is designed to allow occasional large jumps, perhaps away from local modes, thereby improving the chances that the Markov chain will explore the posterior distribution space. Hence, this adaptive scheme allows for further speeding convergence and achieving desirable mixing properties in MCMC chains.

To illustrate this adaptive algorithm in the block-at-a-time MH sampler, I rewrite the notation of the parameter blocks as $\boldsymbol{\Theta} = (\boldsymbol{\theta}_{1,\tau}, \boldsymbol{\theta}_{2,\tau}, ..., \boldsymbol{\theta}_{k+1,\tau})$, where $\boldsymbol{\theta}_{1,\tau} = \boldsymbol{P}$ and $\boldsymbol{\theta}_{j,\tau}$ denotes the parameters in the $(j-1)$th regime for $j = 2, ..., k+1$. And, let $\boldsymbol{\Theta}_{-j}$ denote the vector $\boldsymbol{\Theta}$ excluding the block $\boldsymbol{\theta}_{j,\tau}$. Starting at $g = 1$ with $\boldsymbol{\Theta}^{[1]} = \left(\boldsymbol{\theta}_{1,\tau}^{[1]}, ..., \boldsymbol{\theta}_{k+1,\tau}^{[1]}\right)$, the G_1 random walk MH iterations for $\boldsymbol{\Theta}$ proceed as follows

Step 1. Increment g by 1 and set $\boldsymbol{\Theta}^{[g]}$ equal to $\boldsymbol{\Theta}^{[g-1]}$.

Step 2. For $i = 1, ..., k+1$ in turn, generate $\boldsymbol{\theta}_{i,\tau}^{c}$ as

$$\boldsymbol{\theta}_{i,\tau}^{c} = \boldsymbol{\theta}_{i,\tau}^{[g]} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim \rho N\left(\mathbf{0}, diag\left\{b_i\right\}\right) + (1-\rho) N\left(\mathbf{0}, \omega diag\left\{b_i\right\}\right)$$

and replace $\boldsymbol{\theta}_{i,\tau}^{[g]}$ in $\boldsymbol{\Theta}^{[g]}$ by $\boldsymbol{\theta}_{i,\tau}^{c}$ with the probability $min\left(\zeta_i, 1\right)$, where

$$\zeta_i = \frac{L\left(\boldsymbol{y}_t | \boldsymbol{\theta}_{i,\tau}^{c}, \boldsymbol{\Theta}_{-i}^{[g]}\right) \pi\left(\boldsymbol{\theta}_{i,\tau}^{c}, \boldsymbol{\Theta}_{-i}^{[g]}\right)}{L\left(\boldsymbol{y}_t | \boldsymbol{\Theta}^{[g]}\right) \pi\left(\boldsymbol{\Theta}^{[g]}\right)}$$

Step 3. If $g < G_1$, go to Step 1.

Upon completion, these first G_1 iterations yield the burn-in sample. Following Chen et al. (2012), I set $\rho = 0.95$, $\omega = 100$, and tune the positive number b_i so that the empirical acceptance rate lies in the range $(0.2, 0.45)$ for the ith block. Tuning is done every 100 iterations by increasing b_i when the acceptance rate in the last 100 iterations is higher than 0.45, or decreasing b_i when that rate is lower than 0.2.

At the end of the first G_1 iterations, the burn-in sample mean $\boldsymbol{\mu}_{i,\tau}$ and covariance matrix $\boldsymbol{\Sigma}_{i,\tau}$ of $\boldsymbol{\theta}_{i,\tau}$ with corresponding lower triangular Cholesky factor $\boldsymbol{\Sigma}_{i,\tau}^{1/2}$ are computed for $i = 1, ..., k+1$. The MCMC sampling scheme then continues for G_2 additional iterations according to the following independent kernel MH steps:

Step 4. Increment g by 1 and set $\boldsymbol{\Theta}^{[g]}$ equal to $\boldsymbol{\Theta}^{[g-1]}$.

Step 5. For $i = 1, ..., k+1$ in turn, generate $\boldsymbol{\theta}_{i,\tau}^{c}$ as

$$\boldsymbol{\theta}_{i,\tau}^{c} = \boldsymbol{\mu}_{i,\tau} + \boldsymbol{\Sigma}_{i,\tau}^{1/2} \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim \rho N\left(\mathbf{0}, \boldsymbol{I}\right) + (1-\rho) N\left(\mathbf{0}, \omega \boldsymbol{I}\right)$$

and replace $\boldsymbol{\theta}_{i,\tau}^{[g]}$ in $\boldsymbol{\Theta}^{[g]}$ by $\boldsymbol{\theta}_{i,\tau}^{c}$ with the probability $min\left(\zeta_i, 1\right)$, where

$$\zeta_i = \frac{L\left(\boldsymbol{y}_t|\boldsymbol{\theta}_{i,\tau}^{c}, \boldsymbol{\Theta}_{-i}^{[g]}\right)\pi\left(\boldsymbol{\theta}_{i,\tau}^{c}, \boldsymbol{\Theta}_{-i}^{[g]}\right)q\left(\boldsymbol{\theta}_{i,\tau}^{[g]}\right)}{L\left(\boldsymbol{y}_t|\boldsymbol{\Theta}^{[g]}\right)\pi\left(\boldsymbol{\Theta}^{[g]}\right)q\left(\boldsymbol{\theta}_{i,\tau}^{c}\right)}$$

$$\begin{aligned} q\left(\boldsymbol{\theta}_{i,\tau}\right) \quad \propto \quad & \rho exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}_{i,\tau}-\boldsymbol{\mu}_{i,\tau}\right)'\boldsymbol{\Sigma}_{i,\tau}^{-1}\left(\boldsymbol{\theta}_{i,\tau}-\boldsymbol{\mu}_{i,\tau}\right)\right\} \\ & +\frac{1-\rho}{\omega^{dim(\boldsymbol{\theta}_{i,\tau})/2}} exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}_{i,\tau}-\boldsymbol{\mu}_{i,\tau}\right)'\boldsymbol{\Sigma}_{i,\tau}^{-1}\left(\boldsymbol{\theta}_{i,\tau}-\boldsymbol{\mu}_{i,\tau}\right)\right\} \end{aligned}$$

Step 6. If $g < G_1 + G_2$, go to Step 4.

Observe that the use of $\boldsymbol{\Sigma}_{i,\tau}$ in Step 5 accounts for the posterior correlation among the elements of $\boldsymbol{\theta}_{i,\tau}$, thereby improving the efficiency of the Markov chain. The parameter updates are sequentially repeated until the convergence of the Markov chain is achieved. The burn-in draws are discarded, and the steps are iterated a large number of times to generate draws from which the desired features (means, variances, quantiles, etc.) of the posterior distribution can be estimated consistently.

In this paper, $G_1 = 50,000$ for the random walk MH sampler and $G_2 = 50,000$ with a thinning of 5 for the independent kernel MH sampler, resulting in posterior samples comprising 10,000 draws. The convergence of the IK-MH Markov chains is assessed using the Geweke (1992) test. For each parameter, I also assess the accuracy of its posterior mean by computing the numerical standard error (NSE) according to the batch-means method (Ripley, 1987). In all simulated and real data examples of this paper, it is observed that MCMC chains are well converged inside 50,000 iterations.

5 Simulation

This section carries on a simulation study. In MSQAR nonlinear settings where the number of parameters increases with the number of regimes, it is very convenient to choose parsimonious models that require a low number of parameters. For simplicity in the exposition,

data are simulated from the true model with 2 regimes and autoregressive order 1 as

$$y_t = \begin{cases} 2.0 + 0.2y_{t-1} + 0.5\varepsilon_t, & s_t = 1 \\ -2.0 + 0.4y_{t-1} + \varepsilon_t, & s_t = 2 \end{cases}$$

The true parameter values are referenced based on empirical data estimations in next section. Three underlying distributions are considered for error terms, including a standard normal distribution ($N(0,1)$), a standardized student-t distribution with 3 degrees of freedom (t_3), and a mixed distribution between $N(0,1)$ when $s_t = 1$ and t_3 when $s_t = 2$.

The theoretical τth conditional quantile of y_t can be expressed in a MSQAR form as

$$Q_{y_t}(\tau|\mathbf{y}_{t-1};\theta_{s_t}) = \begin{cases} \theta_{10}(\tau) + \theta_{11}(\tau)y_{t-1}, & s_t = 1 \\ \theta_{20}(\tau) + \theta_{21}(\tau)y_{t-1}, & s_t = 2 \end{cases}$$

with the corresponding quantile parameters as $\theta_{10}(\tau) = 2.0 + 0.5Q_{\varepsilon_t}(\tau)$, $\theta_{11}(\tau) = 0.2$, $\theta_{20}(\tau) = -2.0 + Q_{\varepsilon_t}(\tau)$, and $\theta_{21}(\tau) = 0.4$. $Q_{\varepsilon_t}(\tau)$ is the theoretical τth quantile of a underlying distribution.

200 data replications are simulated for each underlying distribution. 50,00 observations are generated for each data replication but only the last 500 observations are kept for estimation in order to reduce initial effects. MSQAR models are examined in different sample sizes, $T = \{200, 500\}$ and quantile levels, $\tau = \{0.05, 0.25, 0.5, 0.75, 0.95\}$.

Table 1 reports the simulation results. This table includes the true quantile parameters (True), posterior means (PM), standard errors (Std), the root of mean squared errors (RMSE), and the mean absolute deviation (MAD). RMSE and MAD errors in Table 1 are small over different quantile levels and distributions. The small difference between the true and estimated parameters indicates the reasonable accuracy in model estimation. The small standard errors also show a favorable precision in model estimation. Furthermore, the accuracy and the precision of model estimations are improved with the increase in sample sizes considered due to the reduction in RMSEs, MADs and standard errors. As expected,

the model estimation for the less extreme quantiles present smaller RMSE and MAD errors than extreme quantiles. The MSQAR model estimation also shows reasonable performance for the data generated from mixtures of normal and student-t distributions.

[Table 1 about here]

Figure 1 plotting the posterior kernel densities of parameter estimates along with true parameters indicated by the vertical lines. Figure 1 shows that the posteriors well contain the true quantile parameters with a slightly better performance for $\tau = 0.5$. In many cases, the posteriors appear skewed but still with most of the density concentrated near the true parameter values. To save space, Figure 1 plots results for $\tau = 0.05, 0.5, 0.95$ and $N = 200$ from the normal distribution. Other results are similar and available upon request.

[Figure 1 about here]

Following Guerin and Marcellino (2013), Table 2 reports the quadratic probability scores (QPS), absolute probability scores (APS) and log probability scores (LPS) for the quantile autoregressive models with Markov-switching features to check how well these models can estimate the true regimes. QPS, APS and LPS criteria evaluate the qualitative prediction abilities of MSQAR models, that is, to what extent the true quantile regimes are predicted. The predictability of regime 2 is computed for QPS, APS, and LPS as follows:

$$QPS = \frac{2}{T}\sum_{t=1}^{T}\left(\xi_{2,t|t} - I\left(\tilde{s}_t = 2\right)\right)^2 \tag{5.1}$$

$$APS = \frac{1}{T}\sum_{t=1}^{T}\left|\xi_{2,t|t} - I\left(\tilde{s}_t = 2\right)\right| \tag{5.2}$$

$$LPS = -\frac{1}{T}\sum_{t=1}^{T}\left(1 - I\left(\tilde{s}_t = 2\right)\right) log\left(1 - \xi_{2,t|t}\right) + I\left(\tilde{s}_t = 2\right) log\left(\xi_{2,t|t}\right) \tag{5.3}$$

where $\xi_{2,t|t}$ is obtained from (3.2) and $\tilde{s}_t$ is the simulated states. A score of 0 occurs when perfect predictions are made. Note that QPS is bounded between 0 and 2. The worst

score is 2 for QPS and occurs if at each period probability predictions of 0 or 1 are made but turn out to be wrong each time. Note that correct predictions have individual scores between 0 and 0.5, whereas incorrect predictions have individual scores between 0.5 and 2.0 for QPS. Nonetheless, a few incorrect predictions can therefore dominate a majority of correct predictions in QPS scores. For this reason, a modified version of probability scores, absolute probability score (APS), is also considered. Like QPS, the best possible score for APS is 0. The worst score is 1. Here correct predictions have individual scores between 0 and 0.5, whereas incorrect forecasts carry scores between 0.5 and 1. The range for LPS is 0 to ∞. LPS penalizes large prediction errors more than QPS and APS. See also Christoffersen et al. (2007).

[Table 2 about here]

Table 2 shows that all QPS and APS scores are small and less than 0.5, which indicate the dramatic model predictability for switching regimes. The probability scores are slightly lower with the increase in sample sizes. The results also show that regime predictions for lower tails are better than for upper tails. In addition, the statistics of LPS are also smaller than 0.5 which imply that no prediction outliers are penalized.

6 Empirical Applications

Many studies have employed quantile autoregressive models to estimate risks of financial markets and assets. In macroeconomics literature, asymmetric dynamics have also been found for macroeconomic variables. This section applies the proposed MSQAR model to S&P 500 returns for market risk assessment and to real U.S. GDP growth rates (RGDP) and real exchange rates of U.S. dolloars (trade-weighted by major currencies, RTWER) for asymmetric persistence. Monthly and weekly S&P 500 index returns are taken from the Center for Research in Security Prices (CRSP). The quarterly RGDP and RTWER data are taken from Federal Reserve Bank of St. Louis as percent changes from year ago. The data summary in Table 3 show negative skewness for S&P 500 returns and real exchange

rates. The skewness for real GDP growth is positive but small. S&P 500 returns appear to have excess kurtosis. Jarque-Bera tests reject the null of data normality for S&P 500 returns, whereas the tests do not reject the null for real exchange rates. The normality for real GDP is rejected at 10% level. Figure 2 plots the time series of the empirical data.

[Table 3 about here]

[Figure 2 about here]

As discussed in section 5, for empirical illustration, this paper estimates MSQAR of order 1 with 2 regimes to keep a parsimonious parameter space. This paper defines that regime 2 represents more extreme outcomes than regime 1. For instance, at lower tails, quantiles of regime 2 should be more negative or farther into the left tail areas than those of regime 1, which is mostly associated with the periods of economic recessions and crises. In contrary, at upper tails, quantiles of regime 2 should be more positive or farther into the right tail areas than those of regime 1.

6.1 Stock Market Risk

Table 4 reports the estimation results for monthly and weekly S&P 500 returns. The entries are the posterior means of parameters with associated numerical standard errors in parentheses. In general, the values of the Geweke (1992) test statistic in square brackets indicate convergence of the Markov chain to stationarity. Table 5 shows that the numerical standard errors are small and the Markov chain appears to be converged well as indicated by the generally insignificant values of the Geweke (1992) test statistic. Figure 3 also plots the estimated parameters over quantiles with the 5% and 95% intervals of posterior distributions. As seen, the quantile intercepts monotonically increase with the increase of quantile levels. The quantile autoregressive coefficients are close to zero around median, while they deviate from zero at lower and upper tails. The zero coefficients around median seem to suggest market efficiency for S&P 500 index. However, it appears to be less efficient at tails. Moreover, the autoregressive coefficients of regime 2 are larger in magnitude than

those of regime 1. This result implies that markets are less efficient when extreme events occur or during economic recessions and crises. Interestingly, the positive autoregressive coefficients at lower tails suggest that risk expectation is positively impacted by past risks, while the negative autoregressive coefficients at upper tails indicate that during market good times investors is expecting higher risk in future. These results clearly show countercyclical behaviors in financial markets estimated by MSQAR models.

[Table 4 about here]

[Figure 3 about here]

The results also show that the variation of transition probabilities across quantiles is much smaller in regime 1 than in regime 2. The transition probabilities of regime 1 are ranging from 0.85 to 0.985, compared to the range for regime 2 from 0.381 to 0.945. It seems that the more extreme the quantile level is, the lower the persistence of regime 2 (p_{22}) is.[3] Despite the large variation in regime persistence, the unconditional probabilities are very similar across quantiles, i.e., π_1 and π_2 are around 0.84 and 0.16 for each quantile level, respectively.[4] This result is reasonable in that economic conditions provide the common economic states to different parts of a data distribution. However, persistence is possibly varying across quantiles. This observation is further consolidated by Figure 4 plotting the smoothed transition probability $\xi_{s_t=2,t|T}$ for $\tau = 0.05, 0.5$. The shaded areas are NBER-dated business cycles. This figure shows that the fluctuation within each economic recession period is much larger in $\tau = 0.05$ than in $\tau = 0.5$. The responses of the 5% lower tail to the economic recessions are much stronger than those of median, by showing much higher probabilities of switching to regime 2.

[Figure 4 about here]

[3]The regime persistence for regime 1 and 2can be computed as $1/(1-p_{11})$ and $1/(1-p_{22})$, respectively.

[4]Unconditional probabilities of π_1 and π_2 can be obtained as $(1-p_{22})/(2-p_{11}-p_{22})$ and $(1-p_{11})/(2-p_{11}-p_{22})$, respectively.

Value-at-Risk is implicitly defined on quantiles as a one-to-one function of a quantile, over a given time interval, of a conditional return distribution (see Jorion (2000)). For assessing S&P 500 return risks, Figure 5 plots 5% Value-at-Risk (VaR) estimated from the dynamic quantile of $\tau = 0.05$ as $Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t; \hat{\Theta}) = \sum_{i \in S} Q_{y_t}\left(\tau|\boldsymbol{y}_{t-1}, s_t = i; \hat{\theta}_i\right) Pr\left(s_t = i|\boldsymbol{y}_t; \hat{\Theta}\right)$. The dark lines in Figure 5 are the estimated 5% VaR dynamics ($Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t; \hat{\Theta})$) and the top and bottom light lines are the estimated 5% VaR dynamics of regime 1 ($Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t = 1; \hat{\theta}_1)$) and regime 2 ($Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t = 2; \hat{\theta}_2)$), respectively. As seen, the dynamics in regime 2 is larger than in regime 1 due to the larger autoregressive coefficients. This result indicates that market efficiency is different across regimes.

[Figure 5 about here]

The usefulness of the proposed MSQAR model can be immediately recognized from Figure 5. Value-at-Risk estimated from existing methods are undistinguished from different distributions associated with i.e., good times or economic recessions. Thus, the VaR values from those approaches are at best the results of averaging on different economic states. However, Figure 5 shows VaR values for both regime 1 implied by good economic periods and regime 2 associated with economic recessions. Risk states identified by the MSQAR model are particularly beneficial for risk management, as a risk manager would care more about the most extreme scenarios or the worst possible outcomes. For example, to stress-testing a hypothetically stressed financial institution, one should use VaR values estimated from regime 2 ($Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t = 2; \hat{\theta}_2)$) as the worst scenario hypothetically occurring. This may be an appropriate approach to measure systemic risks for considering capital buffer requirement on financial institutions from the perspectives of central banks.

6.2 Asymmetric Persistence in Macroeconomic Dynamics

To study asymmetric dynamics of macroeconomic variables, this paper estimates MSQAR models for percentiles. Table 6 reports the estimation results for real GDP growth rates and real trade-weighted exchange rates. The results of ADF, KPSS, Phillips-Perron and

Zivot-Andrews tests (not reported here) reject the null hypothesis of unit roots for these macroeconomic variables.The entries are the posterior means of parameters with associated numerical standard errors in parentheses and the Geweke (1992) test statistic in square brackets. Table 5 shows that the numerical standard errors are small and the Markov chain appears to be converged well as indicated by the generally insignificant values of the Geweke (1992) test statistic. Figure 6 also plots the estimated parameters over quantile levels with the 5% and 95% intervals of posterior distributions.

[Table 5 about here]

[Figure 6 about here]

Figure 6 shows that the quantile autoregressive coefficients of real GDP growth rates vary over different quantiles, displaying asymmetric dynamics. Upper tails appear to have higher dynamic persistence than lower tails. The quantile autoregressive coefficients of regime 2 has the range from 0.623 to 0.979, compared to the range of 0.779 to 0.874 for the coefficients of regime 1. This result indicates that economic regimes demonstrate different asymmetric dynamics. By contrast, the evidence of asymmetric persistence in real trade-weighted exchange rates is weak due to much less variation across the quantile autoregressive coefficients. This result is consistent with Jarque-Bera test in Table 4 showing that the null of data normality is not rejected for real trade-weighted exchange rates, where it is rejected for real GDP growth rates at 10% confidence level.

In addition, Figure 6 also shows that transition probabilities slightly vary across quantiles in both regimes. It implies that regime persistence of macroeconomic variables is mainly driven by common economic conditions, and hence much less dependent on τ. This result is very different from the regime behaviors of financial markets in section 5, but consistent with the fact that macroeconomic variables are common economic states and factors in an economy.

Table 6 examines the regime predictions of real GDP growth rates. This table reports QPS, APS and LPS values for regime 2 by using NBER-dated business cycles as true

regimes. The probability scores of QPS and APS are smaller than 0.5 across quantiles, which indicates a significant predictability of economic regimes based on real GDP growth. Interestingly, the predictability of regimes from lower tails is much stronger than from upper tails. In addition, LPS values are larger than one at upper tails than at lower tails. This result implies the issue of regime predictive outliers.

[Table 6 about here]

The different regime predictabilities across quantiles are also shown by Figure 7 plotting the smoothed transition probability $\xi_{s_t=2,t|T}$ for $\tau = 0.1, 0.5, 0.9$. The shaded areas are NBER-dated business cycles. As seen, the predicated regimes from $\tau = 0.05$ and $\tau = 0.5$ seem closely to trace NBER dated business cycles, whereas the predicated regimes from $\tau = 0.9$ appear to be lagged. In addition, the responses of the 10% quantile to the economic recessions are much stronger than those of median, by showing much higher probabilities (close to 1) of switching to regime 2. These results suggest that lower tails of real GDP growth rates reveal more information of economic states than upper tails. This might be due to the economic behaviors of risk aversion and also reflect the effects of macroeconomic policies.

[Figure 7 about here]

6.3 Quantile Monotonicity

It is important to evaluate the model by the monotonicity requirement on the conditional quantile functions. If the monotonicity is satisfied, there should be no crossings over quantiles. Severe crossing problems violate the theorem of equivariance to monotone transformation from (1.1) to (1.2). Figure 8 plots the estimated quantiles of each single regime. The straight lines are $Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t = 1; \hat{\theta}_1)$ and $Q_{y_t}(\tau|\boldsymbol{y}_{t-1}, s_t = 2; \hat{\theta}_2)$ for regime 1 and 2. The dots are the scatter plots with y_t as y-axis and y_{t-1} as x-axis. Despite that the MSQAR model is nonlinear, it takes a linear form within a single regime. Quantiles within a regime are not parallel due to its location-scale quantile autoregressive model, unlike location-shift

quantile functions. Quantiles in regime 2 have no crossing issues, while crossing problems occur in regime 1 between $\tau = 0.4, 0.5, 0.6$. Nonetheless, the proportion of violations of the monotonicity in regime 1 is below 2% between $\tau = 0.4, 0.5, 0.6$, except around 10% for the quantiles of regime 1 of real GDP growth rates crossing between $\tau = 0.5$ and $\tau = 0.6$. Overall, Figure 8 does not show severe crossing issues for the data considered in this paper.

[Figure 8 about here]

7 Conclusion

This paper proposes a new location-scale quantile autoregression, so-called Markov-switching quantile autoregression, to characterize behaviors of different parts of a nonstationary time series distribution. The new method directly inferences and estimates dynamic quantiles by allowing the location and scale parameters subject to regime shifts. Unobservable economic regimes are inferred by standard Hamilton filter approach in which quantile error terms follow a three parameter asymmetric Laplace distribution. Bayesian estimation is adopted to deal with some serious computational challenges in this nonlinear model which has differentiable likelihood functions.

The empirical application to S&P 500 returns is able to show countercyclical risk accumulations in financial markets. It also illustrates that the dynamic quantiles associated with economic recessions should be an appropriate extreme scenario for stress-testing hypothetically stressed financial institutions from the perspective of central banks. Furthermore, the estimation results for macroeconomic variables show evidence of asymmetric dynamics for quarterly real GDP growth rates but not for quarterly real trade-weighted U.S. dollars. The transition probabilities are similar across quantiles within a single regime for macroeconomic variables, whereas they vary in financial markets. In addition, this paper has found that the lower tails of real GDP growth provide more valuable information than the upper tails for predicting economic regimes.

References

[1] Ausin, M.C. and H.F. Lopes (2010) Time-varying joint distribution through copulas. Computational Statistics and Data Analysis 54: 2383-2399

[2] Bauwens, L., A. Preminger and J.V.K. Rombouts (2010) Theory and inference for a Markov-Switching GARCH model. The Econometrics Journal 13: 218-244

[3] Cai, Y. and J. Stander (2008) Quantile Self-Exciting Threshold Autoregressive Time Series Models. Journal of Time Series Analysis 29(1): 186-202

[4] Cai, Y. (2010) Forecasting for quantile self-exciting threshold autoregressive time series models. Biometrika 97(1): 199-208

[5] Chen, Q., R. Gerlach, and Z. Lu (2012) Bayesian Value-at-Risk and expected shortfall forecasting via the asymmetric Laplace distribution. Computational Statistics and Data Analysis 56(11): 3498-3516

[6] Chernozhukov, V. and H. Hong (2003) An MCMC approach to classical estimation. Journal of Econometrics 115: 293-346

[7] Cheung, Y. and U.G. Erlandsson (2005) Exchange rates and Markov-Switching dynamics. Journal of Business & Economic Statistics 23(3): 314-320

[8] Chib, S. and E. Greenberg (1995) Understanding the Metropolis-Hastings algorithm. American Statistician 49: 327-335

[9] Christoffersen, P.F., F.X. Diebold, R.S. Mariano, A.S. Tay and Y.K. Tse (2007) Direction-of-change forecasts based on conditional variance, skewness and kurtosis dynamics: international evidence. Journal of Financial Forecasting 1(2): 1-22

[10] Gerlach, R., 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(4): 481-492

[11] Geweke, J. and H. Tanizaki (2001) Bayesian estimation of state-space models using the Metropolis-Hastings algorithm within Gibbs sampling. Computational Statistics & Data Analysis 37: 151-170

[12] Geweke, J. (1992) Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernardo, J., Berger, J., David, A., Smith, A. (Eds.), Bayesian Statistics. Vol. 4. Oxford University Press, Oxford, pp. 169{193.

[13] Gray, S.F. (1996) Modeling the conditional distribution of interest rates as a regime-switching process. Journal of Financial Economics 42: 27-62

[14] Guerin, P. and M. Marcellino (2013) Markov-Switching MIDAS Models. Journal of Business & Economic Statistics 31(1): 45-56

[15] Guidolin, M. (2012) Markov Switching Models in Empirical Finance. Advances in Econometrics, ISBN: 978-1-78052-526-6

[16] Hamilton, J.D. (1994) Time Series Analysis. Princeton University Press

[17] Hamilton, J., D. Waggoner and T. Zha (2007) Normalization in econometrics. Econometric Reviews 26: 221-252

[18] Hamilton, J.D. and R. Susmel (1994) Autoregressive conditional heteroskedasticity and changes in regime. Journal of Econometrics 64: 307-333

[19] Jorion, P. (2000) Value-at-Risk: The New Benchmark for Managing Financial Risk. McGraw-Hill. ISBN-13: 978-0071355025

[20] Kim, C.J. (1994) Dynamic linear models with Markov-switching. Journal of Econometrics 60: 1-22

[21] Kim, C.J., J. Piger, and R. Startz (2008) Estimation of Markov regime-switching regression models with endogenous switching. Journal of Econometrics 143: 263-273

[22] Koenker, R. (2005) Quantile Regression. Cambridge University Press.

[23] Koenker, R. and G. Bassett (1978) Regression quantile. Econometrica 46: 33-50

[24] Koenker, R. and Z. Xiao (2006) Quantile Autoregression. Journal of the American Statistical Association 101(475): 980-990

[25] Ripley, B. (1987) Stochastic Simulation. John Wiley, New York

[26] Sims, C.A. and T. Zha (2006) Were There Regime Switching in US Monetary Policy. American Economic Review 96(1): 54-81

[27] Tierney, L. (1994) Markov Chains for Exploring Posterior Distributions. Ann. Statist. 22: 1701-1728

[28] Vrontos, I., P. Dellaportas, D. Politis (2002) Full Bayesian inference for GARCH and EGARCH models. Journal of Business and Economic Statistics 18: 187-198

[29] Yu, K. and J. Zhang (2005) A Three-Parameter Asymmetric Laplace Distribution and Its Extension. Communications in Statistics- Theory and Methods 34: 1867-1879

[30] Yu, K. and R.A. Moyeed (2001) Bayesian quantile regression. Statistics & Probability Letters 54: 437-447

Table 1: Simulation Results

	(1) Normal errors							
	True	$N = 200$				$N = 500$		
		PM	Std	RMSE	MAD	PM	Std	RMSE	MAD
$\tau = 0.05$								
p_{11}	0.900	0.883	0.034	0.042	0.031	0.895	0.019	0.022	0.017
p_{22}	0.900	0.890	0.034	0.039	0.030	0.901	0.022	0.025	0.019
$\theta_{10}(\tau)$	1.178	1.262	0.113	0.120	0.099	1.249	0.081	0.092	0.075
$\theta_{11}(\tau)$	0.200	0.181	0.038	0.210	0.168	0.184	0.032	0.178	0.145
$\theta_{20}(\tau)$	-3.645	-3.584	0.295	0.082	0.065	-3.606	0.193	0.054	0.043
$\theta_{21}(\tau)$	0.400	0.401	0.097	0.172	0.200	0.400	0.059	0.147	0.116
$\tau = 0.25$									
p_{11}	0.900	0.887	0.034	0.040	0.029	0.898	0.019	0.021	0.016
p_{22}	0.900	0.889	0.032	0.038	0.029	0.899	0.021	0.024	0.018
$\theta_{10}(\tau)$	1.663	1.671	0.088	0.053	0.041	1.665	0.060	0.036	0.029
$\theta_{11}(\tau)$	0.200	0.194	0.030	0.151	0.122	0.195	0.024	0.120	0.099
$\theta_{20}(\tau)$	-2.674	-2.677	0.221	0.083	0.067	-2.660	0.123	0.046	0.037
$\theta_{21}(\tau)$	0.400	0.394	0.067	0.168	0.133	0.402	0.039	0.098	0.081
$\tau = 0.5$									
p_{11}	0.900	0.889	0.034	0.040	0.029	0.900	0.019	0.021	0.015
p_{22}	0.900	0.888	0.032	0.038	0.029	0.897	0.021	0.023	0.018
$\theta_{10}(\tau)$	2.000	1.997	0.085	0.043	0.035	1.988	0.055	0.028	0.022
$\theta_{11}(\tau)$	0.200	0.196	0.028	0.142	0.112	0.198	0.021	0.106	0.085
$\theta_{20}(\tau)$	-2.000	-2.057	0.209	0.108	0.089	-2.041	0.119	0.063	0.052
$\theta_{21}(\tau)$	0.400	0.382	0.064	0.165	0.139	0.388	0.037	0.097	0.080
$\tau = 0.75$									
p_{11}	0.900	0.891	0.035	0.040	0.030	0.903	0.019	0.022	0.018
p_{22}	0.900	0.884	0.032	0.040	0.031	0.894	0.021	0.024	0.019
$\theta_{10}(\tau)$	2.337	2.323	0.087	0.038	0.031	2.311	0.057	0.027	0.021
$\theta_{11}(\tau)$	0.200	0.203	0.029	0.144	0.116	0.206	0.021	0.107	0.054
$\theta_{20}(\tau)$	-1.326	-1.499	0.231	0.118	0.172	-1.469	0.148	0.105	0.125
$\theta_{21}(\tau)$	0.400	0.358	0.075	0.177	0.173	0.365	0.047	0.140	0.115
$\tau = 0.95$									
p_{11}	0.900	0.884	0.040	0.048	0.035	0.896	0.024	0.027	0.021
p_{22}	0.900	0.863	0.036	0.057	0.046	0.874	0.025	0.040	0.033
$\theta_{10}(\tau)$	2.822	2.779	0.123	0.046	0.037	2.772	0.080	0.034	0.027
$\theta_{11}(\tau)$	0.200	0.217	0.042	0.225	0.188	0.215	0.033	0.181	0.148
$\theta_{20}(\tau)$	-0.355	-0.453	0.124	0.144	0.380	-0.451	0.103	0.123	0.322
$\theta_{21}(\tau)$	0.400	0.337	0.086	0.187	0.215	0.328	0.067	0.145	0.204

RMSE and MAD are the root of mean squared errors and the mean absolute deviation errors.

	(2) t_3 errors							
	True	$N = 200$				$N = 500$		
		PM	Std	RMSE	MAD	PM	Std	RMSE	MAD
$\tau = 0.05$								
p_{11}	0.900	0.879	0.031	0.042	0.033	0.885	0.021	0.029	0.022
p_{22}	0.900	0.880	0.041	0.050	0.037	0.892	0.024	0.028	0.022
$\theta_{10}(\tau)$	1.321	1.428	0.101	0.111	0.094	1.441	0.077	0.108	0.094
$\theta_{11}(\tau)$	0.200	0.191	0.035	0.180	0.148	0.193	0.028	0.142	0.117
$\theta_{20}(\tau)$	-3.359	-3.319	0.398	0.119	0.093	-3.319	0.246	0.074	0.059
$\theta_{21}(\tau)$	0.400	0.416	0.118	0.197	0.240	0.407	0.080	0.121	0.161
$\tau = 0.25$								
p_{11}	0.900	0.885	0.031	0.038	0.029	0.891	0.021	0.026	0.020
p_{22}	0.900	0.879	0.035	0.045	0.033	0.889	0.021	0.027	0.021
$\theta_{10}(\tau)$	1.779	1.770	0.062	0.035	0.028	1.777	0.043	0.024	0.018
$\theta_{11}(\tau)$	0.200	0.202	0.024	0.118	0.095	0.202	0.016	0.078	0.060
$\theta_{20}(\tau)$	-2.442	-2.465	0.156	0.064	0.049	-2.443	0.091	0.037	0.031
$\theta_{21}(\tau)$	0.400	0.393	0.048	0.120	0.093	0.400	0.029	0.073	0.058
$\tau = 0.5$								
p_{11}	0.900	0.886	0.031	0.037	0.028	0.892	0.021	0.025	0.020
p_{22}	0.900	0.876	0.034	0.046	0.034	0.886	0.021	0.028	0.022
$\theta_{10}(\tau)$	2.000	1.994	0.058	0.029	0.022	1.993	0.034	0.017	0.014
$\theta_{11}(\tau)$	0.200	0.199	0.023	0.115	0.091	0.201	0.013	0.067	0.051
$\theta_{20}(\tau)$	-2.000	-2.025	0.117	0.060	0.048	-2.025	0.074	0.039	0.028
$\theta_{21}(\tau)$	0.400	0.394	0.037	0.095	0.077	0.396	0.024	0.062	0.049
$\tau = 0.75$								
p_{11}	0.900	0.886	0.032	0.039	0.029	0.892	0.022	0.026	0.020
p_{22}	0.900	0.870	0.034	0.050	0.039	0.881	0.022	0.033	0.026
$\theta_{10}(\tau)$	2.221	2.218	0.075	0.034	00.025	2.212	0.043	0.020	0.016
$\theta_{11}(\tau)$	0.200	0.199	0.027	0.133	0.101	0.202	0.016	0.080	0.063
$\theta_{20}(\tau)$	-1.558	-1.643	0.137	0.103	0.083	-1.647	0.085	0.079	0.065
$\theta_{21}(\tau)$	0.400	0.383	0.042	0.113	0.092	0.385	0.027	0.076	0.062
$\tau = 0.95$								
p_{11}	0.900	0.881	0.039	0.048	0.036	0.888	0.026	0.032	0.024
p_{22}	0.900	0.853	0.037	0.067	0.055	0.861	0.025	0.051	0.044
$\theta_{10}(\tau)$	2.679	2.690	0.186	0.069	0.053	2.675	0.128	0.048	0.038
$\theta_{11}(\tau)$	0.200	0.197	0.066	0.153	0.265	0.202	0.047	0.137	0.186
$\theta_{20}(\tau)$	-0.641	-0.596	0.110	0.185	0.141	-0.596	0.073	0.134	0.106
$\theta_{21}(\tau)$	0.400	0.354	0.066	0.201	0.167	0.358	0.048	0.158	0.130

RMSE and MAD are the root of mean squared errors and the mean absolute deviation errors.

	(3) Mixed errors								
	True	$N = 200$				$N = 500$			
		PM	Std	RMSE	MAD	PM	Std	RMSE	MAD
$\tau = 0.05$									
p_{11}	0.900	0.885	0.032	0.042	0.030	0.896	0.019	0.025	0.017
p_{22}	0.900	0.893	0.032	0.039	0.028	0.905	0.020	0.023	0.018
$\theta_{10}(\tau)$	1.178	1.271	0.118	0.128	0.105	1.252	0.085	0.096	0.078
$\theta_{11}(\tau)$	0.200	0.182	0.038	0.212	0.167	0.186	0.029	0.160	0.126
$\theta_{20}(\tau)$	-3.359	-3.406	0.226	0.127	0.096	-3.348	0.155	0.076	0.059
$\theta_{21}(\tau)$	0.400	0.416	0.128	0.163	0.163	0.409	0.085	0.114	0.169
$\tau = 0.25$									
p_{11}	0.900	0.886	0.034	0.041	0.030	0.895	0.021	0.023	0.016
p_{22}	0.900	0.887	0.031	0.038	0.027	0.898	0.019	0.022	0.016
$\theta_{10}(\tau)$	1.663	1.671	0.082	0.049	0.040	1.668	0.059	0.035	0.029
$\theta_{11}(\tau)$	0.200	0.197	0.033	0.166	0.13	0.199	0.022	0.108	0.086
$\theta_{20}(\tau)$	-2.442	-2.462	0.153	0.063	0.049	-2.446	0.094	0.039	0.030
$\theta_{21}(\tau)$	0.400	0.400	0.046	0.115	0.090	0.400	0.029	0.072	0.057
$\tau = 0.5$									
p_{11}	0.900	0.886	0.035	0.039	0.029	0.894	0.021	0.022	0.015
p_{22}	0.900	0.883	0.030	0.037	0.026	0.893	0.019	0.021	0.016
$\theta_{10}(\tau)$	2.000	1.998	0.080	0.040	0.033	1.990	0.053	0.027	0.022
$\theta_{11}(\tau)$	0.200	0.200	0.031	0.153	0.125	0.202	0.022	0.111	0.088
$\theta_{20}(\tau)$	-2.000	-2.037	0.119	0.062	0.047	-2.033	0.076	0.041	0.033
$\theta_{21}(\tau)$	0.400	0.393	0.037	0.093	0.073	0.393	0.024	0.062	0.049
$\tau = 0.75$									
p_{11}	0.900	0.886	0.035	0.042	0.032	0.894	0.023	0.026	0.020
p_{22}	0.900	0.878	0.030	0.041	0.031	0.888	0.019	0.025	0.019
$\theta_{10}(\tau)$	2.337	2.329	0.084	0.036	0.028	2.315	0.061	0.028	0.022
$\theta_{11}(\tau)$	0.200	0.202	0.033	0.167	0.134	0.208	0.025	0.129	0.105
$\theta_{20}(\tau)$	-1.558	-1.648	0.130	0.104	0.081	-1.651	0.087	0.082	0.070
$\theta_{21}(\tau)$	0.400	0.382	0.040	0.110	0.088	0.382	0.026	0.079	0.064
$\tau = 0.95$									
p_{11}	0.900	0.877	0.041	0.052	0.039	0.884	0.027	0.035	0.027
p_{22}	0.900	0.856	0.034	0.062	0.050	0.866	0.022	0.045	0.039
$\theta_{10}(\tau)$	2.822	2.813	0.153	0.054	0.042	2.798	0.095	0.035	0.028
$\theta_{11}(\tau)$	0.200	0.204	0.058	0.218	0.167	0.208	0.035	0.181	0.145
$\theta_{20}(\tau)$	-0.641	-0.600	0.103	0.173	0.135	-0.601	0.087	0.128	0.098
$\theta_{21}(\tau)$	0.400	0.355	0.062	0.192	0.160	0.354	0.044	0.158	0.133

RMSE and MAD are the root of mean squared errors and the mean absolute deviation errors.

Table 2: Summary Statistics for the Predictability of Simulated Regimes.

	Normal			t_3			Mixed		
	QPS	APS	LPS	QPS	APS	LPS	QPS	APS	LPS
$N = 200$									
$\tau = 0.05$	0.019	0.025	0.046	0.049	0.042	0.251	0.043	0.042	0.206
$\tau = 0.25$	0.013	0.014	0.025	0.018	0.015	0.040	0.013	0.011	0.031
$\tau = 0.5$	0.006	0.011	0.014	0.014	0.011	0.038	0.010	0.010	0.029
$\tau = 0.75$	0.033	0.032	0.065	0.034	0.026	0.101	0.030	0.025	0.088
$\tau = 0.95$	0.055	0.052	0.138	0.056	0.050	0.305	0.048	0.044	0.286
$N = 500$									
$\tau = 0.05$	0.019	0.025	0.046	0.045	0.038	0.196	0.038	0.038	0.136
$\tau = 0.25$	0.012	0.013	0.024	0.016	0.012	0.038	0.012	0.011	0.030
$\tau = 0.5$	0.006	0.011	0.013	0.014	0.011	0.040	0.010	0.010	0.025
$\tau = 0.75$	0.031	0.031	0.063	0.033	0.025	0.102	0.030	0.024	0.086
$\tau = 0.95$	0.052	0.050	0.140	0.054	0.048	0.297	0.046	0.042	0.274

QPS, APS and LPS represent quadratic, absolute and log probability scores, respectively.

Table 3: Data Summary Statistics

	Monthly S&P 500	Weekly S&P 500	Real GDP	Real TWER
Sample Periods	1926:01-2013:02	01/09/1950-02/25/2013	1948Q1-2013Q2	1974Q1-2013Q2
# of obs.	1047	3294	263	159
Mean	0.461	0.136	3.263	-0.132
Median	0.907	0.282	3.200	0.201
Std. dev.	5.505	2.091	2.675	7.087
Skewness	-0.525	-0.567	0.004	-0.256
Kurtosis	10.75	8.744	0.692	-0.044
Jarque-Bera	<0.001	<0.001	0.058	0.412

Note: the $p-$values are reported for Jarque-Bera statistics test of data normality.

Table 4: Estimation Results for S&P 500 Index Returns

	Monthly S&P 500					Weekly S&P 500				
	$\tau = 0.05$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.95$	$\tau = 0.05$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.95$
p_{11}	0.902	0.925	0.985	0.987	0.842	0.850	0.868	0.985	0.947	0.787
	(0.018)	(0.055)	(0.010)	(0.009)	(0.016)	(0.017)	(0.034)	(0.011)	(0.102)	(0.019)
	[1.440]	[-0.447]	[-0.405]	[0.062]	[-0.274]	[0.631]	[0.065]	[1.273]	[0.685]	[-1.832]
p_{22}	0.572	0.625	0.914	0.917	0.381	0.389	0.641	0.945	0.880	0.415
	(0.045)	(0.063)	(0.043)	(0.039)	(0.066)	(0.046)	(0.037)	(0.027)	(0.146)	(0.044)
	[-0.685]	[-0.570]	[-2.405]	[-0.483]	[-0.244]	[-0.276]	[0.446]	[-0.803]	[0.568]	[-0.509]
$\theta_{10}(\tau)$	-4.121	-0.114	1.045	3.522	4.971	-1.619	-0.039	0.346	1.066	1.647
	(0.089)	(0.051)	(0.064)	(0.062)	(0.063)	(0.039)	(0.041)	(0.040)	(0.132)	(0.035)
	[-2.444]	[-0.464]	[0.648]	[-0.108]	[0.836]	[-0.322]	[-0.747]	[-1.678]	[0.723]	[-1.007]
$\theta_{11}(\tau)$	0.084	0.035	0.026	-0.087	-0.082	0.085	0.015	0.035	-0.025	-0.032
	(0.042)	(0.026)	(0.027)	(0.031)	(0.042)	(0.036)	(0.026)	(0.022)	(0.081)	(0.020)
	[1.835]	[-0.307]	[-0.518]	[0.098]	[0.064]	[-0.328]	[1.246]	[-0.788]	[-0.616]	[-0.126]
$\theta_{20}(\tau)$	-15.77	-3.978	-2.124	5.784	12.05	-5.865	-1.952	-0.094	2.422	4.851
	(0.147)	(0.047)	(0.318)	(0.174)	(0.200)	(0.094)	(0.040)	(0.147)	(0.147)	(0.080)
	[-0.478]	[0.101]	[-1.615]	[-0.212]	[0.169]	[-0.296]	[1.404]	[0.965]	[-0.835]	[-0.665]
$\theta_{21}(\tau)$	0.378	0.115	0.085	0.042	0.032	0.207	0.010	0.016	-0.196	-0.251
	(0.039)	(0.045)	(0.048)	(0.054)	(0.046)	(0.038)	(0.017)	(0.056)	(0.069)	(0.018)
	[-0.224]	[-0.342]	[0.889]	[0.395]	[0.621]	[0.317]	[0.348]	[0.124]	[0.617]	[0.538]

Values in parentheses are numerical standard errors and the Geweke (1992) test statistic in square brackets.

The test distribution for Geweke (1992) statistic is standard normal distribution.

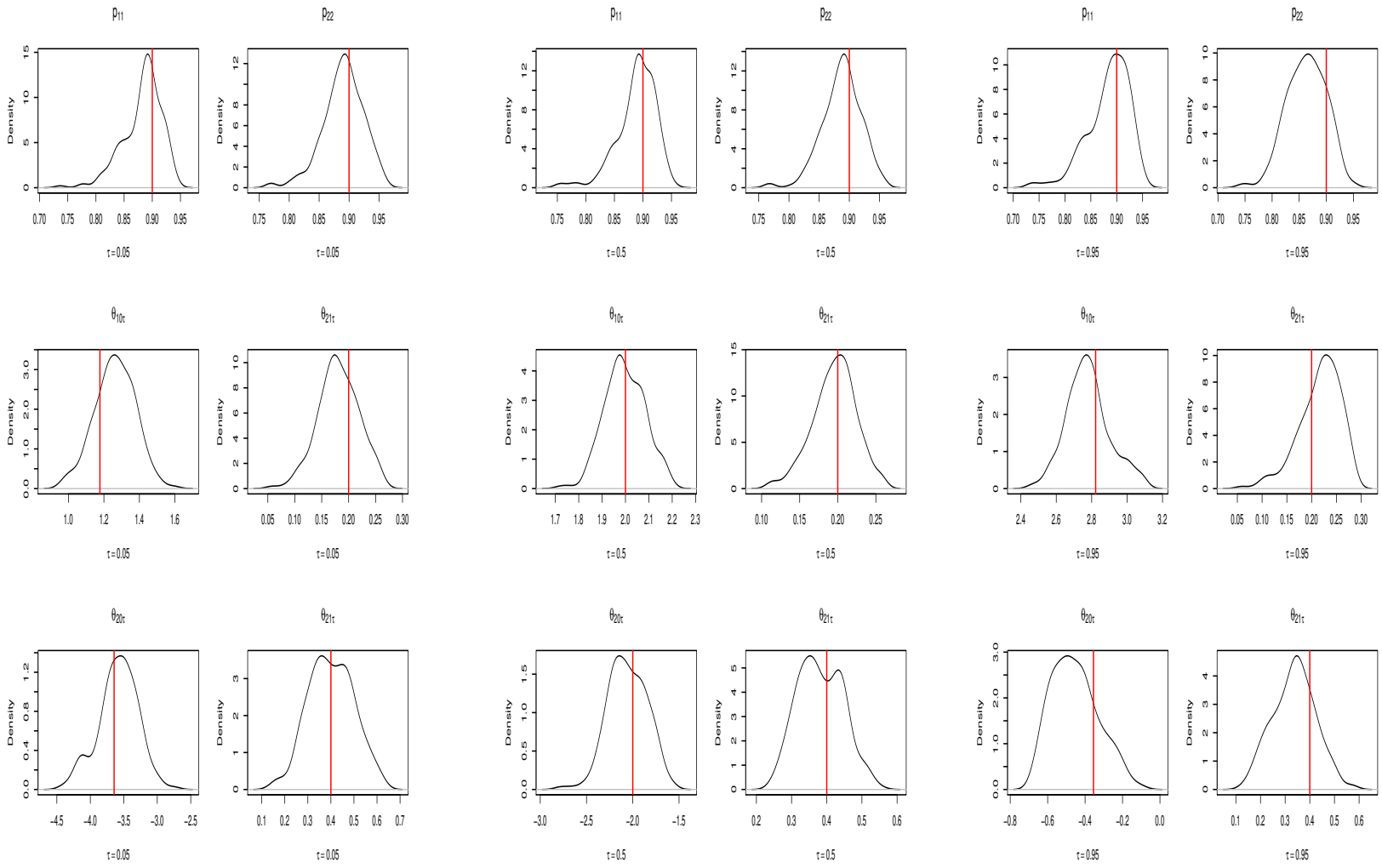
Table 5: Estimation Results for Macroeconomic Variables

	Real GDP Growth Rates						Real Trade-weighted Exchange Rates					
	p_{11}	p_{22}	$\theta_{10}(\tau)$	$\theta_{11}(\tau)$	$\theta_{20}(\tau)$	$\theta_{21}(\tau)$	p_{11}	p_{22}	$\theta_{10}(\tau)$	$\theta_{11}(\tau)$	$\theta_{20}(\tau)$	$\theta_{21}(\tau)$
$\tau = 0.1$	0.927	0.647	-0.178	0.781	-2.310	0.623	0.789	0.500	-2.318	0.774	-7.949	0.761
	(0.023)	(0.046)	(0.087)	(0.032)	(0.162)	(0.099)	(0.056)	(0.093)	(0.291)	(0.055)	(0.264)	(0.079)
	[-1.085]	[1.193]	[0.699]	[-1.382]	[0.737]	[1.090]	[-1.110]	[0.409]	[1.434]	[-1.067]	[1.083]	[-1.205]
$\tau = 0.2$	0.926	0.667	0.196	0.779	-1.734	0.668	0.811	0.573	-1.149	0.777	-6.199	0.788
	(0.021)	(0.047)	(0.085)	(0.037)	(0.089)	(0.047)	(0.066)	(0.053)	(0.209)	(0.047)	(0.327)	(0.047)
	[0.375]	[-1.077]	[-2.107]	[0.067]	[-1.015]	[-0.535]	[1.368]	[-0.837]	[-1.439]	[1.294]	[-1.110]	[-0.585]
$\tau = 0.3$	0.930	0.638	0.310	0.801	-1.561	0.716	0.832	0.582	-0.620	0.784	-5.156	0.781
	(0.043)	(0.063)	(0.048)	(0.031)	(0.143)	(0.076)	(0.054)	(0.041)	(0.103)	(0.051)	(0.211)	(0.057)
	[0.559]	[0.085]	[0.227]	[-0.407]	[-1.193]	[-1.127]	[-0.392]	[-0.794]	[0.700]	[0.911]	[-0.083]	[0.489]
$\tau = 0.4$	0.943	0.687	0.630	0.784	-1.125	0.758	0.849	0.554	-0.323	0.799	-3.488	0.832
	(0.054)	(0.134)	(0.182)	(0.055)	(0.276)	(0.208)	(0.046)	(0.066)	(0.110)	(0.048)	(0.193)	(0.058)
	[-0.125]	[0.076]	[0.113]	[0.292]	[0.047]	[0.175]	[1.295]	[-1.579]	[-0.804]	[0.771]	[-0.858]	[-0.453]
$\tau = 0.5$	0.939	0.706	0.890	0.779	-0.629	0.883	0.926	0.489	-0.196	0.820	-2.307	0.770
	(0.015)	(0.039)	(0.068)	(0.034)	(0.114)	(0.054)	(0.033)	(0.056)	(0.092)	(0.025)	(0.222)	(0.052)
	[-0.925]	[0.229]	[0.043]	[-0.178]	[1.245]	[-0.825]	[0.457]	[-1.276]	[-1.231]	[0.357]	[-0.514]	[0.460]
$\tau = 0.6$	0.965	0.609	0.917	0.790	3.302	0.938	0.808	0.486	0.485	0.829	2.893	0.831
	(0.025)	(0.043)	(0.032)	(0.016)	(0.222)	(0.043)	(0.066)	(0.065)	(0.130)	(0.058)	(0.299)	(0.126)
	[-1.041]	[-0.643]	[0.874]	[-0.637]	[-0.242]	[-1.456]	[-1.150]	[0.160]	[-1.222]	[-0.654]	[-0.241]	[-0.100]
$\tau = 0.7$	0.955	0.609	0.974	0.809	3.514	0.958	0.792	0.465	0.995	0.810	4.660	0.878
	(0.018)	(0.039)	(0.036)	(0.020)	(0.057)	(0.013)	(0.094)	(0.089)	(0.177)	(0.053)	(0.419)	(0.135)
	[-1.217]	[0.589]	[0.051]	[-0.745]	[0.441]	[-0.345]	[0.666]	[-2.952]	[-0.132]	[-0.109]	[1.301]	[1.115]
$\tau = 0.8$	0.940	0.592	1.113	0.834	3.626	0.971	0.779	0.440	1.676	0.791	5.963	0.900
	(0.023)	(0.054)	(0.042)	(0.020)	(0.074)	(0.015)	(0.073)	(0.088)	(0.230)	(0.074)	(0.303)	(0.096)
	[1.427]	[-0.311]	[0.032]	[-0.465]	[0.421]	[-0.301]	[0.370]	[1.522]	[-0.580]	[1.120]	[0.095]	[-0.253]
$\tau = 0.9$	0.924	0.579	1.375	0.874	3.762	0.979	0.808	0.412	2.799	0.754	8.335	0.820
	(0.029)	(0.058)	(0.065)	(0.027)	(0.175)	(0.013)	(0.086)	(0.069)	(0.226)	(0.033)	(0.506)	(0.102)
	[0.437]	[-0.124]	[-0.288]	[1.182]	[0.423]	[1.072]	[-1.057]	[-1.162]	[-0.114]	[0.029]	[-1.054]	[0.991]

Table 6: Real GDP Growth Rates: Predicability of Regime 2

	QPS	APS	LPS
$\tau = 0.1$	0.135	0.111	0.281
$\tau = 0.2$	0.131	0.117	0.235
$\tau = 0.3$	0.134	0.119	0.223
$\tau = 0.4$	0.132	0.132	0.219
$\tau = 0.5$	0.151	0.179	0.259
$\tau = 0.6$	0.403	0.227	1.039
$\tau = 0.7$	0.439	0.249	1.075
$\tau = 0.8$	0.481	0.277	1.134
$\tau = 0.9$	0.516	0.298	1.370

Figure 1: Posteriors of the Parameter Estimates with the true parameters indicated by the vertical lines for $\tau = 0.05, 0.5, 0.95$.



(1) $\tau = 0.05$ (2) $\tau = 0.5$ (3) $\tau = 0.95$

Figure 2: Data for Monthly and Weekly S&P 500, Quarterly Real GDP Growth Rates and Real Trade-weighted Exchange Rates

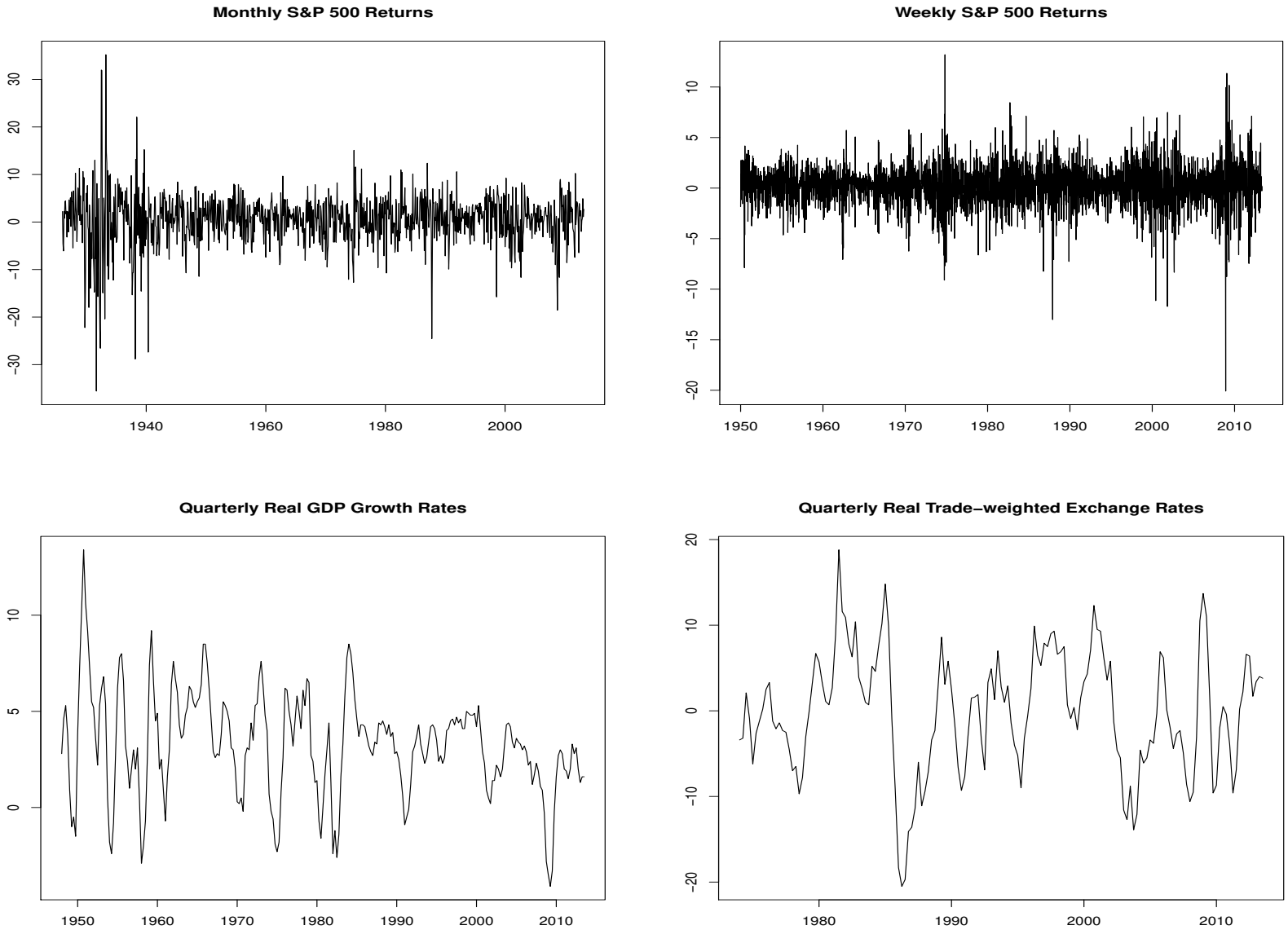
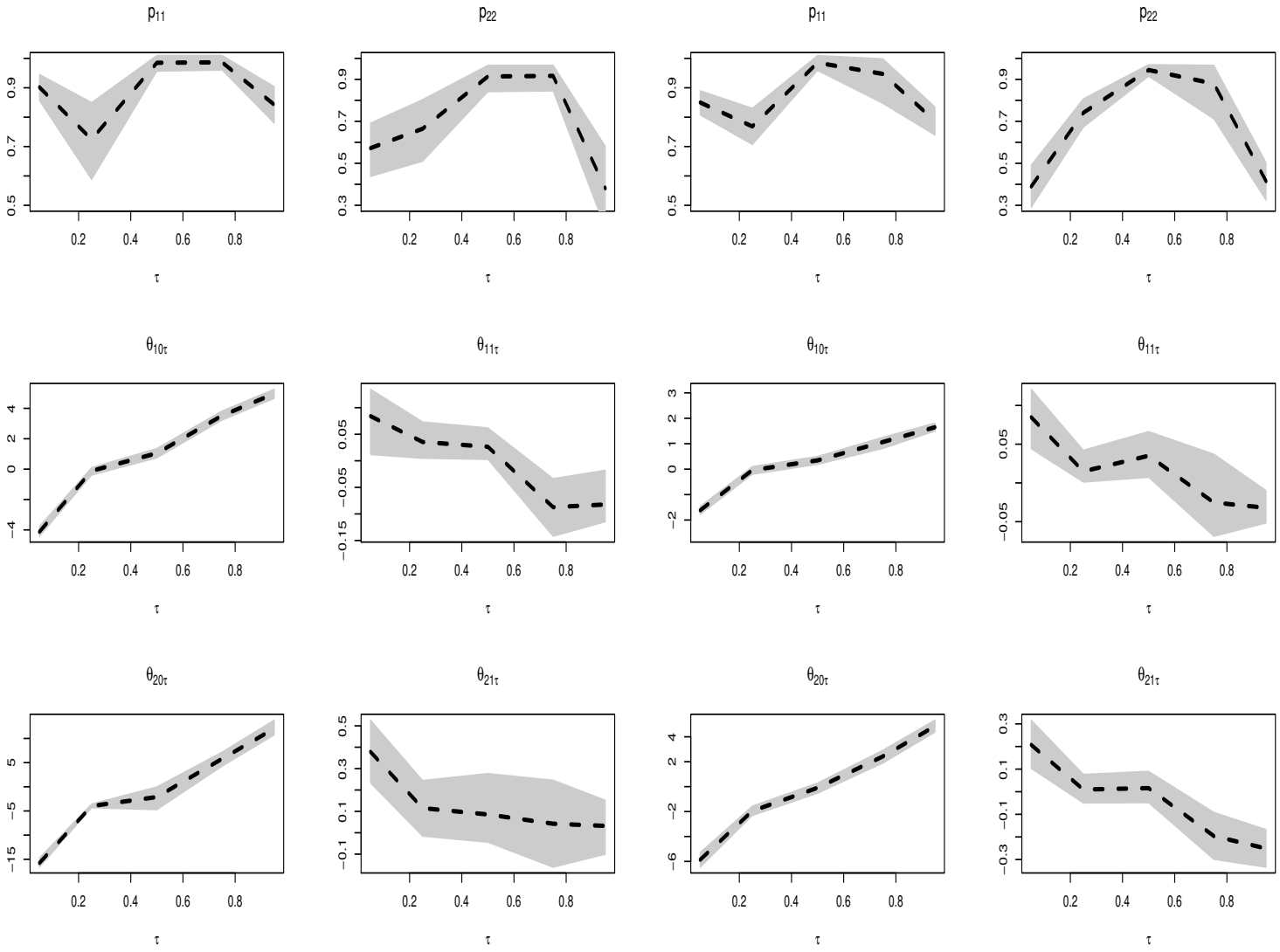


Figure 3: Quantile Parameter Estimation: S&P 500



(1) Monthly S&P 500 Returns

(2) Weekly S&P 500 Returns

Figure 4: Smoothed Transition Probability. The shaded areas are NBER-dated business cycles

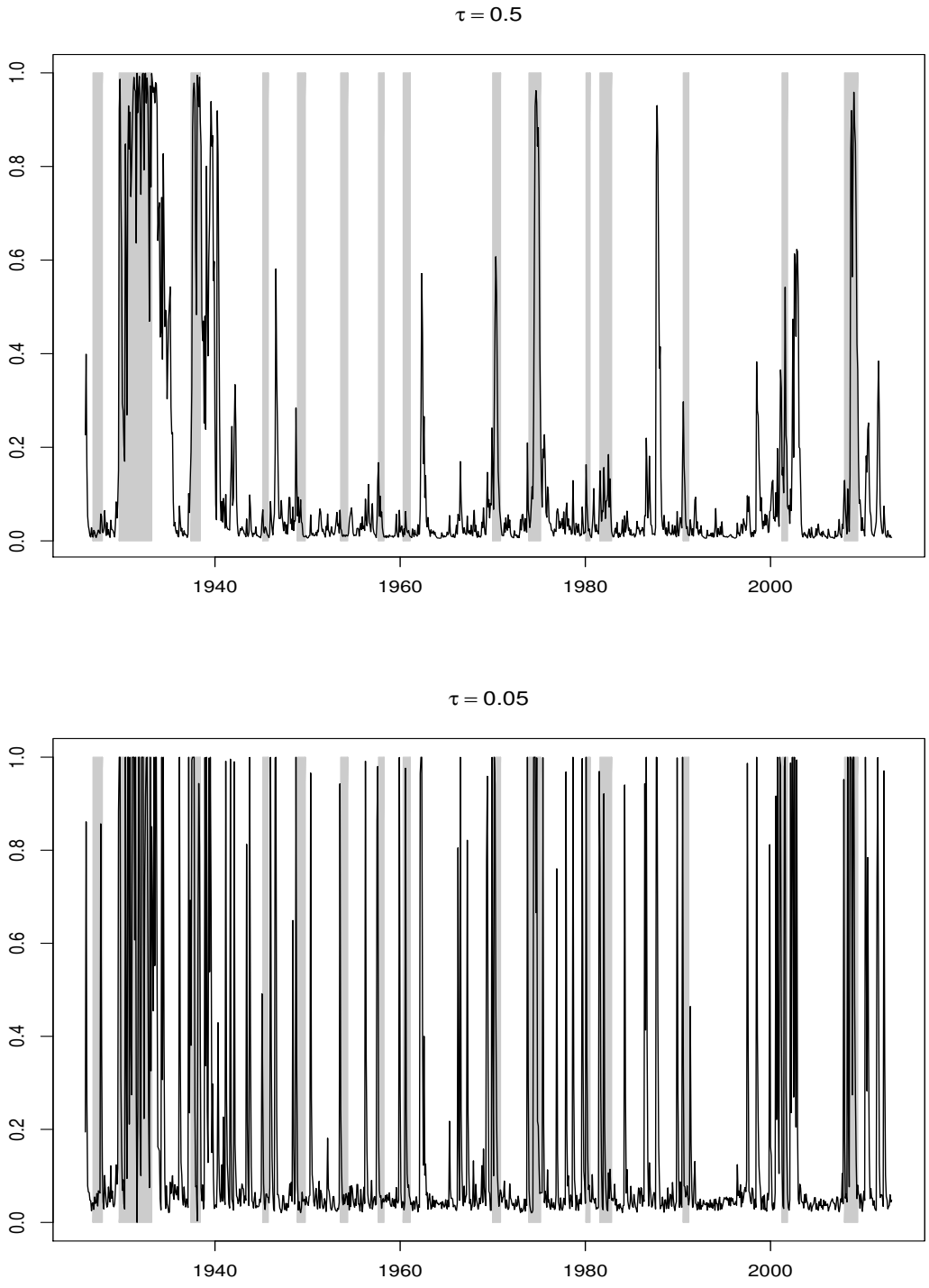


Figure 5: The Estimated Quantiles for $Q_{y_t}(\tau = 0.05|s_t = 1)$ (top light lines), $Q_{y_t}(\tau = 0.05|s_t)$ (dark lines), $Q_{y_t}(\tau = 0.05|s_t = 2)$(bottom light lines)

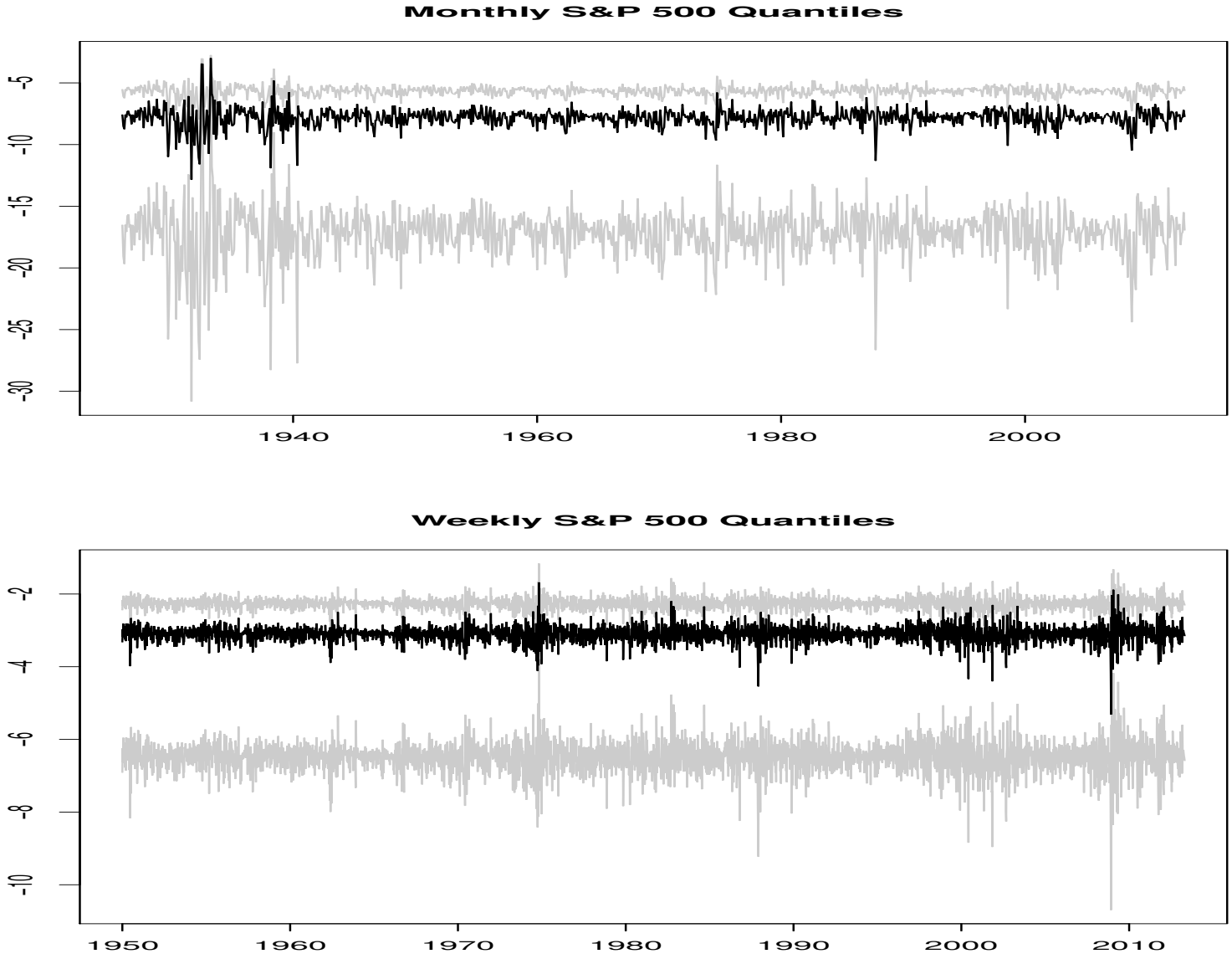
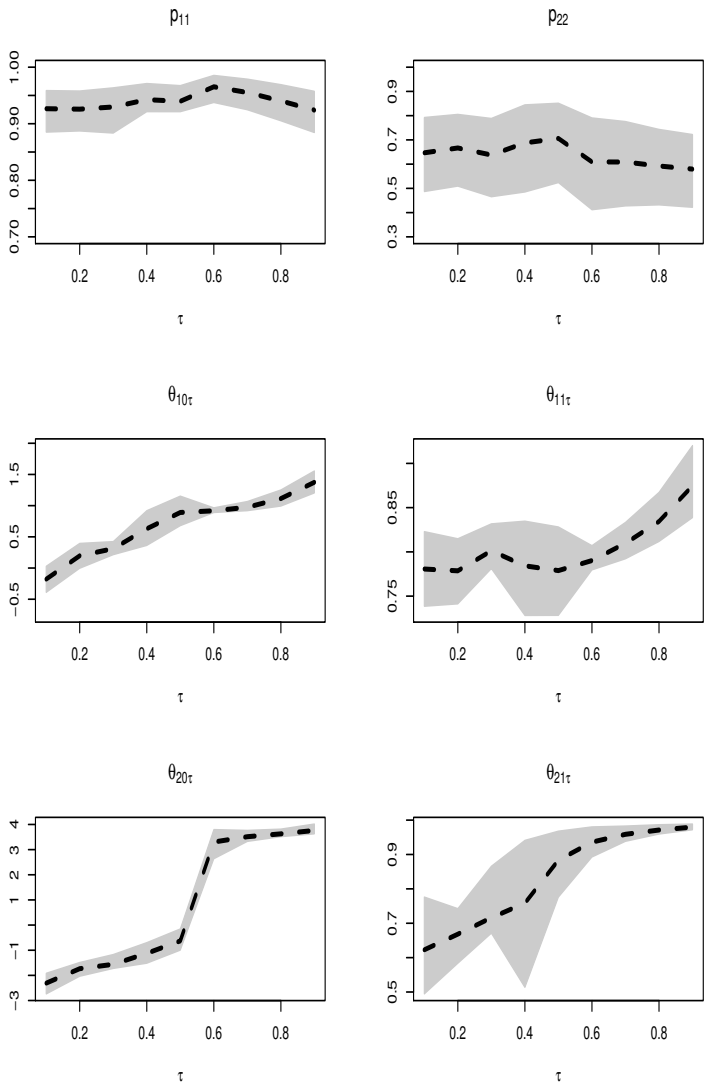
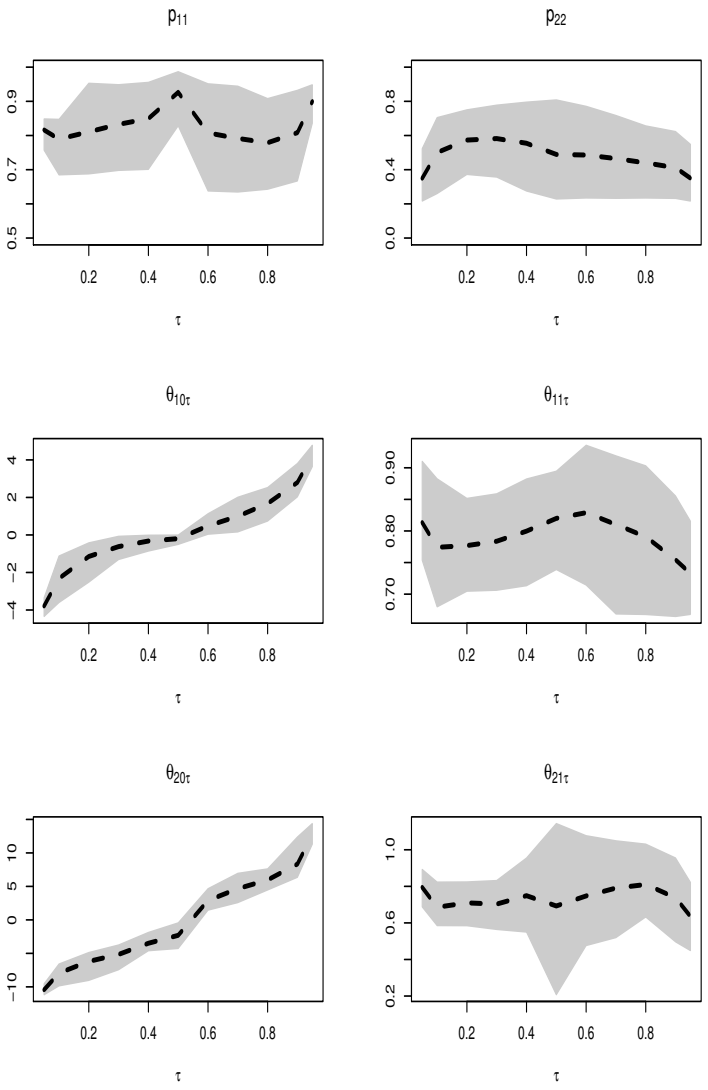


Figure 6: Quantile Parameter Estimations for Macroeconomic Variables



(1) Real GDP Growth Rates



(2) Real Trade-weighted Exchange Rates

Figure 7:
Smoothed Transition Probability for Real GDP. The shaded areas are NBER-dated business cycles

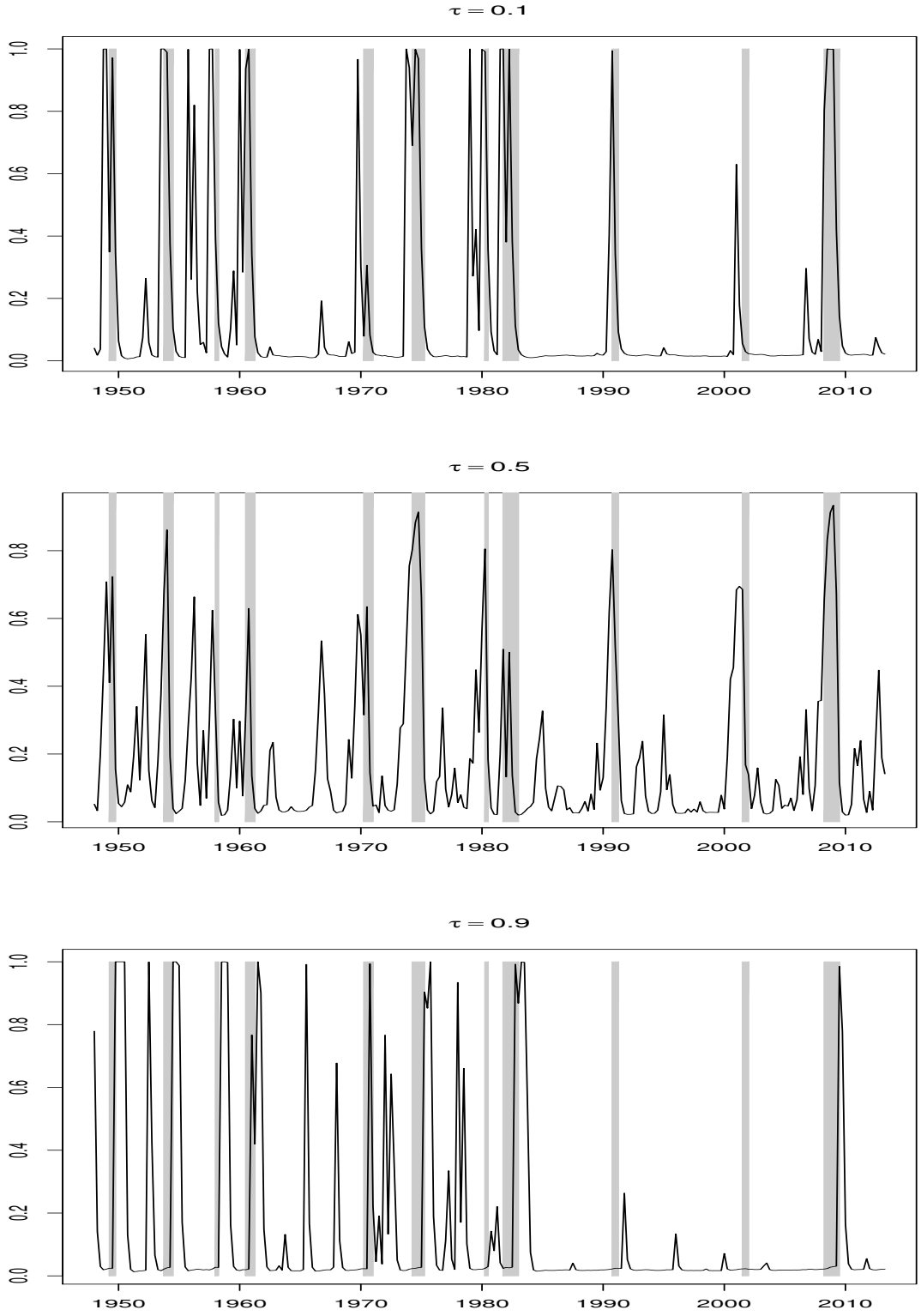
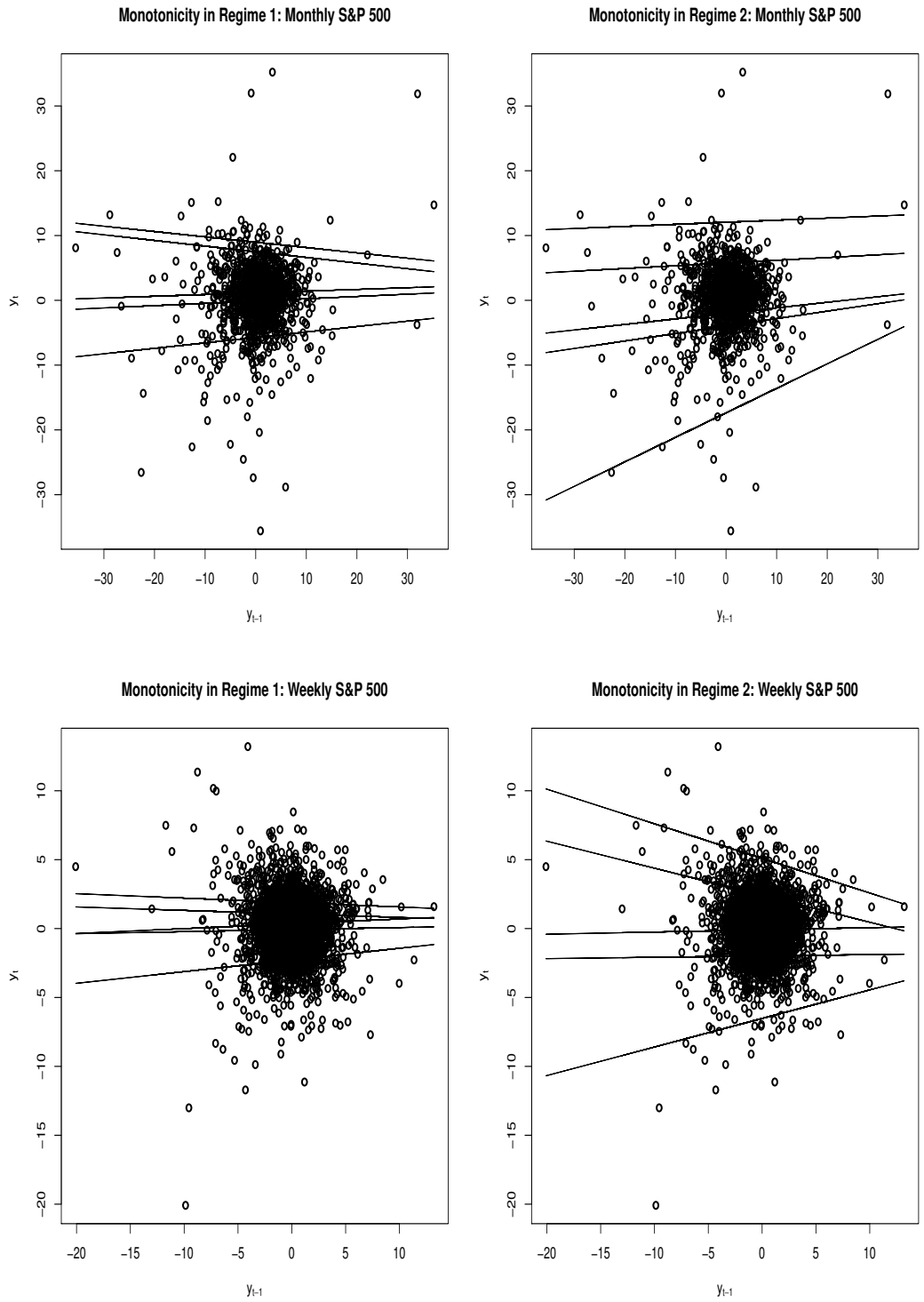
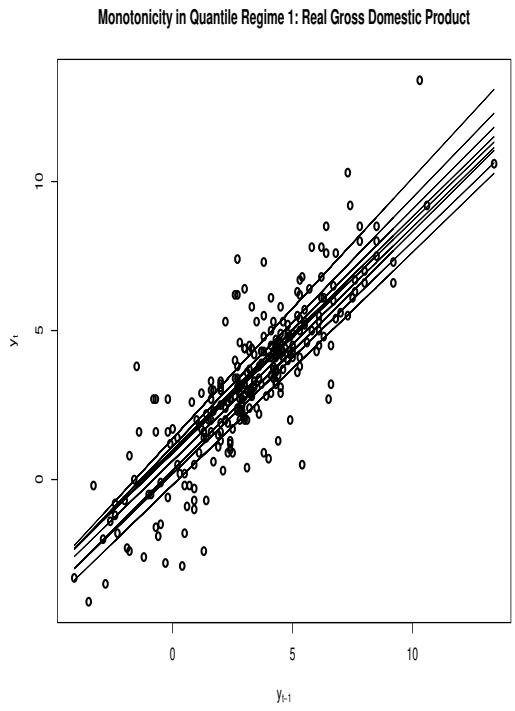
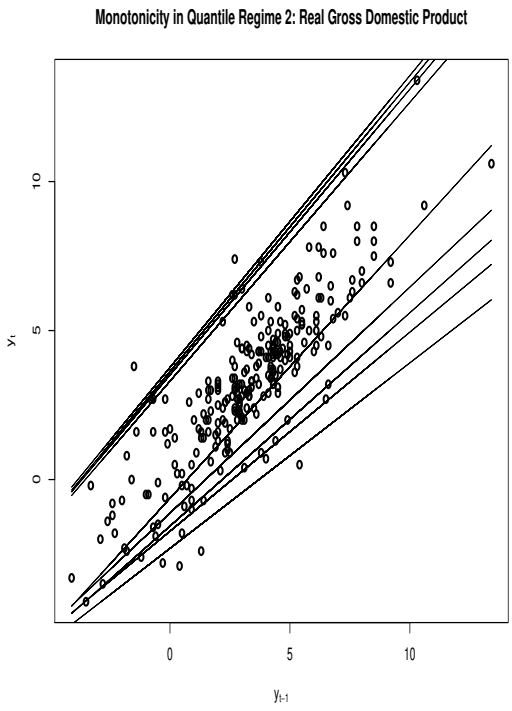
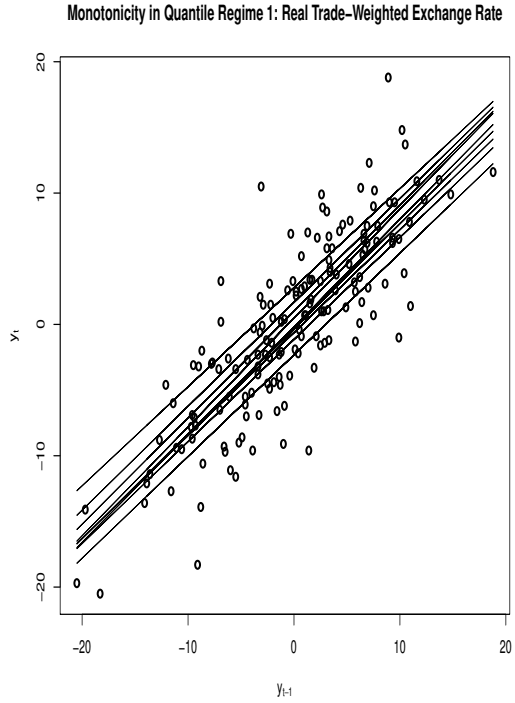


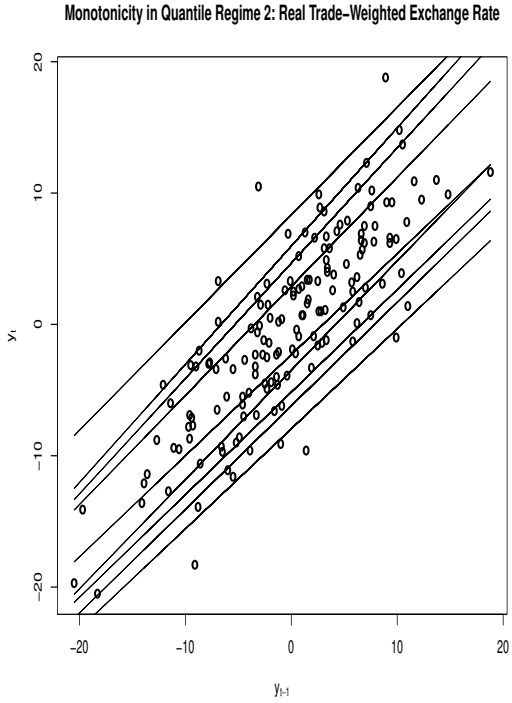
Figure 8: Quantile Monotonicity. This figure plots dynamic quantile for each single regime.




Monotonicity in Quantile Regime 1: Real Gross Domestic Product
y_t
y_{t-1}


Monotonicity in Quantile Regime 2: Real Gross Domestic Product
y_t
y_{t-1}


Monotonicity in Quantile Regime 1: Real Trade-Weighted Exchange Rate
y_t
y_{t-1}


Monotonicity in Quantile Regime 2: Real Trade-Weighted Exchange Rate
y_t
y_{t-1}