Statistica Sinica **30** (2020), 809-827
doi:https://doi.org/10.5705/ss.202017.0324

CONDITIONAL QUANTILE ESTIMATION FOR HYSTERETIC AUTOREGRESSIVE MODELS

Degao Li[1], Ruochen Zeng[2], Liwen Zhang[3], Wai Keung Li[2] and Guodong Li[2]

[1]*Jiaxing University,* [2]*University of Hong Kong,*
and [3]*Shanghai University of Finance and Economics*

Abstract: The phenomenon of hysteresis has been observed in many economic time series, especially in unemployment rates. To study the hysteretic patterns at different quantiles, this study considers a conditional quantile estimation for hysteretic autoregressive models, and derives its asymptotic properties. Simulation experiments are conducted to evaluate the finite-sample performance of our method, and its usefulness is further demonstrated by an analysis of the growth rates of unemployment rates.

Key words and phrases: Autoregression, conditional quantile estimation, hysteretic model, threshold model.

1. Introduction

The threshold model (Tong and Lim (1980)) has been shown to be highly successful in interpreting time-irreversibility, limit cycles, asymmetric dynamics, and so on; see Tong (1990) for a comprehensive exposition. However, it is well known that the model does not work well around the boundaries between regimes (Wu and Chen (2007)), possibly due to a sudden change in the probability structure when a threshold process switches regimes. This problem has been reduced, to some extent, by other regime-switching models, such as the smooth-transition threshold models in Chan and Tong (1986), discrete-state Markov switching models in Hamilton (1989) and McCulloch and Tsay (1994), and a threshold variable-driven switching model in Wu and Chen (2007). However, although they grant the threshold model greater flexibility by changing the piecewise linear structure or introducing latent random variables to the regime-switching mechanism, in general, they lack a physical interpretation.

The phenomenon of hysteresis has been observed in many economic time series, especially in unemployment rates (Brunello (1990); Roed (2002); Song and Wu (1997, 1998)). Economic theory decomposes the unemployment rate into two components: a short-term cyclical and a long-term natural rate. When there is a negative shock to the economy, the cyclical rate will rise, but the natural rate may also rise, owing to the propagation effect of the shock. There are also microeconomic interpretations. First, the unemployed may lose skills, and thus have more difficulty in returning to work. Second, owing to the wage bargaining institution and labor turnover costs, incumbent workers have an incentive to bargain for higher wages when the economy begins to recover, which renders the wage level higher than the market equilibrium level (Blanchard and Summers (1986)). As a result, the observed unemployment rate will be pushed up, and remains high for longer than expected; see Amable et al. (1995) and Perez-Alonso and Sanzo (2011).

Li et al. (2015) proposed a hysteretic autoregressive (HAR) time series model that combines threshold models and the phenomenon of hysteresis,

$$y_t = \begin{cases} \theta_{01} + \theta_{11} y_{t-1} + \cdots + \theta_{p1} y_{t-p} + \sigma_1 \varepsilon_t, & R_t = 1, \\ \theta_{02} + \theta_{12} y_{t-1} + \cdots + \theta_{p2} y_{t-p} + \sigma_2 \varepsilon_t, & R_t = 0, \end{cases} \tag{1.1}$$

with the regime indicator

$$R_t = \begin{cases} 1 & y_{t-d} \le r_L, \\ 0 & y_{t-d} > r_U, \\ R_{t-1} & r_L < y_{t-d} \le r_U. \end{cases} \tag{1.2}$$

Because the hysteresis zone $(r_L, r_U]$ acts as a buffer for the regime switching, a better model fit is expected. Moreover, when $r_L = r_U$, the hysteretic model reduces to a threshold model, which corresponds to the case without hysteresis; see Section 2. Owing to the phenomenon of hysteresis in unemployment rates, it is of interest to apply the hysteretic model in (1.1) and (1.2) to the corresponding sequence. Moreover, the movement of the natural unemployment rate can be approximately separated into two phases, one rising up and the other returning back, which may correspond to the upper and lower regimes for (1.2).

In the meanwhile, since the work of Koenker and Bassett (1978), the quantile regression has become a valuable tool for analyzing the conditional quantile functions of a response variable. Compared with the conditional mean regression, a quantile regression provides a more comprehensive analysis on how predictors may influence different aspects of the conditional distribution of the response; see Koenker (2005) for a comprehensive introduction. Koenker and Xiao (2006) proposed the quantile autoregressive (AR) model, the first quantile model in the literature on time series. Kato (2009) extended the convexity arguments to the scenario under which the estimators are derived as a stochastic process, providing a useful tool to theoretically study quantile regression and threshold models. Kuan, Michalopoulos and Xiao (2017) studied time series models with possible threshold structures at some quantile levels, and Cai and Stander (2008) considered a Bayesian approach to estimate the conditional quantiles of threshold AR processes; see also Cai (2010). Galvao et al. (2014) developed a uniform test for linearity against the threshold effect of the quantile regression, and Zhang, Wang and Zhu (2014) suggested a CUSUM-type test for the threshold at some quantile levels. In particular, Galvao, Montes-Rojas and Olmo (2011) considered a conditional quantile estimation for the threshold AR model, finding that the structures differ significantly by quantile for the US monthly unemployment rate. Together with the phenomenon of hysteresis in economic time series, this motivates us to consider the conditional quantile estimation for HAR models.

The remainder of this paper is organized as follows. Section 2 describes the model setting and the estimating procedure. The asymptotic properties are derived in Section 3. Section 4 conducts simulation experiments to evaluate the finite-sample performance of the conditional quantile estimation. A sequence of growth rates of the unemployment rate is analyzed in Section 5. Section 6 gives a short conclusion and discussion. All technical proofs are provided in the online Supplementary Material.

2. Conditional Quantile Estimation for HAR Models

For a fixed $\tau \in (0,1)$, the τth conditional quantile of the HAR process, generated by (1.1) and (1.2), has the form of

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = \begin{cases} \theta_{01,\tau} + \theta_{11,\tau} y_{t-1} + \cdots + \theta_{p1,\tau} y_{t-p}, & R_{t,\tau} = 1, \\ \theta_{02,\tau} + \theta_{12,\tau} y_{t-1} + \cdots + \theta_{p2,\tau} y_{t-p}, & R_{t,\tau} = 0, \end{cases} \tag{2.1}$$

with the regime indicator

$$R_{t,\tau} = \begin{cases} 1 & y_{t-d_\tau} \le r_{L,\tau}, \\ 0 & y_{t-d_\tau} > r_{U,\tau}, \\ R_{t-1,\tau} & r_{L,\tau} < y_{t-d_\tau} \le r_{U,\tau}, \end{cases}$$

where $\mathcal{F}_t$ is the σ-field generated by $\{y_t, y_{t-1}, \dots\}$, d_τ is the delay parameter, y_{t-d_τ} is the hysteresis variable, and $(r_{L,\tau}, r_{U,\tau}]$ is the hysteresis zone. Here, we use subscripts to emphasize the dependence on τ. After some algebraic calculation, it can be verified that, in the almost surely sense,

$$R_{t,\tau} = I\{y_{t-d_\tau} \le r_{L,\tau}\} + I\{r_{L,\tau} < y_{t-d_\tau} \le r_{U,\tau}\} R_{t-1,\tau} \tag{2.2}$$

$$= I\{y_{t-d_\tau} \le r_{L,\tau}\} + \sum_{j=0}^{\infty}\prod_{i=0}^{j} I\{r_{L,\tau} < y_{t-d_\tau-i} \le r_{U,\tau}\} I\{y_{t-d_\tau-j-1} \le r_{L,\tau}\},$$

where $I(\cdot)$ is the indicator function. When $r_{L,\tau} = r_{U,\tau}$, it holds that $R_{t,\tau} = I\{y_{t-d_\tau} \le r_{L,\tau}\}$, which corresponds to a threshold AR model without hysteresis. Let $n_0 = \max(p, d_{\max})$. For an observed sequence $\{y_t, -n_0+1 \le t \le n\}$, we next consider the conditional quantile estimation for model (2.1).

Denote the parameter vector by $\lambda_\tau = (\theta_\tau^T, d_\tau, r_{L,\tau}, r_{U,\tau})^T$, where $\theta_{1,\tau} = (\theta_{01,\tau}, \ldots, \theta_{p1,\tau})^T$, $\theta_{2,\tau} = (\theta_{02,\tau}, \ldots, \theta_{p2,\tau})^T$, and $\theta_\tau = (\theta_{1,\tau}^T, \theta_{2,\tau}^T)^T$. Let Θ be a compact set of $\mathbb{R}^{2p+2}$, $[a, b]$ be a predetermined interval, and $d_{\max}$ be a predetermined positive integer. For the true value of parameter vector λ_τ, we assume that θ_τ^0 is an interior point of Θ, $a < r_{L,\tau}^0 < r_{U,\tau}^0 < b$, and $d_\tau^0 \in D = \{1, \ldots, d_{\max}\}$.

Let $x_t = (1, y_{t-1}, \ldots, y_{t-p})^T$. Then, model (2.1) can be rewritten in the following compact form:

$$Q_{y_t}(\tau|\mathcal{F}_{t-1}) = x_t^T \theta_{1,\tau} R_t(r_{L,\tau}, r_{U,\tau}, d_\tau) + x_t^T \theta_{2,\tau}[1 - R_t(r_{L,\tau}, r_{U,\tau}, d_\tau)], \qquad (2.3)$$

where, from (2.2), the regime indicator function $R_t(r_{L,\tau}, r_{U,\tau}, d_\tau)$ depends on past observations infinitely far away, because $r_{L,\tau} < r_{U,\tau}$. For fixed $r_{L,\tau}, r_{U,\tau}$, and d_τ, the first few observations of the hysteresis variable, say $\{y_{1-d_\tau}, \ldots, y_{t_0-d_\tau}\}$, may fall within the hysteresis zone $(r_{L,\tau}, r_{U,\tau}]$, such that we cannot identify the corresponding regimes. For simplicity, we can assign them to the lower regime, and denote the resulting regime indicator function by $\widetilde{R}_t(r_{L,\tau}, r_{U,\tau}, d_\tau)$. Note that the exact value of $R_{t_0+1}(r_{L,\tau}, r_{U,\tau}, d_\tau)$ is known, and it holds that $\widetilde{R}_t(r_L, r_{U,\tau}, d_\tau) = R_t(r_{L,\tau}, r_{U,\tau}, d_\tau)$, for $t_0 < t \le n$.

Let $L_n(\lambda_\tau) = \sum_{t=1}^n \rho_\tau[y_t - M_t(\lambda_\tau)]$ be the loss function, where

$$M_t(\lambda_\tau) = x_t^T \theta_{1,\tau} R_t(r_{L,\tau}, r_{U,\tau}, d_\tau) + x_t^T \theta_{2,\tau}[1 - R_t(r_{L,\tau}, r_{U,\tau}, d_\tau)],$$

and $\rho_\tau(u) = u[\tau - I(u < 0)]$ is the check function. When the regime indicator function $R_t(r_{L,\tau}, r_{U,\tau}, d_\tau)$ in $M_t(\lambda_\tau)$ and $L_n(\lambda_\tau)$ is replaced with $\widetilde{R}_t(r_{L,\tau}, r_{U,\tau}, d_\tau)$, we denote them by $\widetilde{M}_t(\lambda_\tau)$ and $\widetilde{L}_n(\lambda_\tau)$, respectively. The conditional quantile estimator of model (2.1) can then be defined as

$$\widehat{\lambda}_{n,\tau} = \operatorname{argmin} \widetilde{L}_n(\lambda_\tau),$$

where $\widehat{\lambda}_{n,\tau} = (\widehat{\theta}_{n,\tau}^T, \widehat{d}_\tau, \widehat{r}_{L,\tau}, \widehat{r}_{U,\tau})^T$.

From (2.3), numerically minimizing $\widetilde{L}_n(\lambda_\tau)$ for each fixed $r_{L,\tau}$, $r_{U,\tau}$, and d_τ is equivalent to performing a linear quantile regression. We denote the resulting minimizer by $\widetilde{\theta}_{n,\tau}(r_{L,\tau}, r_{U,\tau}, d_\tau)$. Note that $\widetilde{L}_n[\widetilde{\theta}_{n,\tau}(r_{L,\tau}, r_{U,\tau}, d_\tau),\, r_{L,\tau}, r_{U,\tau}, d_\tau]$ is a stepwise function with possible jumps at $d_\tau \in D$ and

$$(r_{L,\tau}, r_{U,\tau}) \in \{(y_{t-d_\tau}, y_{s-d_\tau}) : 1 \le t, s \le n; y_{t-d_\tau} \le y_{s-d_\tau}\};$$

see Li and Li (2008, 2011). As a result, it can be minimized by searching over all jumps, and the corresponding minimizer is our conditional quantile estimator $(\widehat{d}_\tau, \widehat{r}_{L,\tau}, \widehat{r}_{U,\tau})$. We can verify that $\widehat{\theta}_{n,\tau} = \widetilde{\theta}_{n,\tau}(\widehat{r}_{L,\tau}, \widehat{r}_{U,\tau}, \widehat{d}_\tau)$.

For the initial value of regime indicator function $R_t(r_{L,\tau}, r_{U,\tau}, d_\tau)$, we may alternatively assign the first t_0 observations to the upper regime, and denote by $\widetilde{R}_t^*(r_{L,\tau}, r_{U,\tau}, d_\tau)$ the resulting regime indicator function. Let $\widetilde{L}_n^*(\lambda_\tau)$ be the corresponding loss function, and $\widehat{\lambda}_{n,\tau}^* = \operatorname{argmin} \widetilde{L}_n^*(\lambda_\tau)$. We can choose $\widehat{\lambda}_{n,\tau}^*$ as the estimator if $\widetilde{L}_n^*(\widehat{\lambda}_{n,\tau}^*) < \widetilde{L}_n(\widehat{\lambda}_{n,\tau})$. In practice, the values of a and b can be set to some percentiles of the observed data.

We next adopt the Bayesian information criterion (BIC) of Lee, Noh and Park (2014) to select the order p. Denote $\widetilde{R}_{t,\tau} = \widetilde{R}_t(\widehat{r}_{L,\tau}, \widehat{r}_{U,\tau}, \widehat{d}_\tau)$, for simplicity. By temporarily assuming that ε_t in (1.1) follows an asymmetric Laplace distribution, with density

$$f(x) = \tau(1-\tau)\exp\{-\rho_\tau(x)\},$$

we can define

$$\mathrm{BIC}(p) = 2n_1 \log \widehat{\sigma}_{1n} + (p+1)\log n_1 + 2n_2 \log \widehat{\sigma}_{2n} + (p+1)\log n_2, \qquad (2.4)$$

where $\widehat{\sigma}_{1n} = n_1^{-1}\sum_{t=1}^n \rho_\tau(y_t - x_t^T\widehat{\theta}_{1n,\tau})\widetilde{R}_{t,\tau}$, $\widehat{\sigma}_{2n} = n_2^{-1}\sum_{t=1}^n \rho_\tau(y_t - x_t^T\widehat{\theta}_{2n,\tau})(1 - \widetilde{R}_{t,\tau})$, $n_1 = \sum_{t=1}^n \widetilde{R}_{t,\tau}$, $n_2 = n - n_1$, and $\widehat{\theta}_{n,\tau} = (\widehat{\theta}_{1n,\tau}^T, \widehat{\theta}_{2n,\tau}^T)^T$. We can define the Akaike information criterion (AIC) in a similar way. Moreover, it is possible to consider different orders, say p_1 and p_2, for the two regimes of model (2.1) in the information criteria proposed above.

3. Asymptotic Results

Assumption 1. *It holds that* $\theta_{1,\tau}^0 \neq \theta_{2,\tau}^0$, $P(y_t \in [a,b]) < 1$, *and the time series* $\{y_t\}$ *is strictly stationary, with* $E(|y_t|^{2+\varsigma}) < \infty$, *for some* $\varsigma > 0$.

Theorem 1. *If Assumption* 1 *holds, then* $\widehat{\lambda}_{n,\tau} \to \lambda_\tau^0$ *almost surely as* $n \to \infty$, *where* $\lambda_\tau^0 = (\theta_\tau^{0T}, r_{L,\tau}^0, r_{U,\tau}^0, d_\tau^0)^T$ *and* $\theta_\tau^0 = (\theta_{1,\tau}^{0T}, \theta_{2,\tau}^{0T})^T$.

Note that the delay parameter d_τ only takes discrete integer values. It then holds that $\widehat{d}_\tau = d_\tau^0$ when the sample size n is sufficiently large. Without loss of generality, the true delay parameter d_τ^0 is assumed to be known for the remainder of this section; hence, we omit it from the parameter vector λ_τ and the corresponding functions.

Let $Y_t = (y_t, \ldots, y_{t-p+1}, R_t)^T$. From (1.1) and (1.2), it can be verified that $\{Y_t\}$ is a Markov chain, and we denote its k-step transition probability by $P^k(x, z)$. Let

$$\Omega_0 = \text{diag}\left\{E\{x_t x_t^T R_{t,\tau}\}, E\{x_t x_t^T (1 - R_{t,\tau})\}\right\},$$
$$\Omega_1 = \text{diag}\left\{E\{f_t[F_t^{-1}(\tau)] x_t x_t^T R_{t,\tau}\}, E\{f_t[F_t^{-1}(\tau)] x_t x_t^T (1 - R_{t,\tau})\}\right\},$$

where $f_t(\cdot)$ and $F_t(\cdot)$ are the density and distribution functions, respectively, of y_t, conditional on $\mathcal{F}_{t-1}$.

Assumption 2. *The Markov chain $\{Y_t\}$ has a unique invariant measure $\pi(\cdot)$, such that $\exists K > 0$ and $\exists \kappa \in [0, 1)$, $\forall x \in \mathbb{R}^p \times \{0, 1\}$ and $\forall k \in \mathbb{N}$, $\|P^k(x, \cdot) - \pi(\cdot)\|_v \le K(1 + \|x\|)\kappa^k$, where $\|\cdot\|_v$ and $\|\cdot\|$ denote the total variation norm and the Euclidean norm, respectively.*

Assumption 3. *There exist $p - 1$ constants $z_{p-1}, \ldots, z_{p-d_\tau+1}, z_{p-d_\tau-1}, \ldots, z_0$, such that $Z^T(\theta_{1,\tau}^0 - \theta_{2,\tau}^0) \neq 0$ for all $z_{p-d_\tau} \in [r_{L,\tau}^0, r_{U,\tau}^0]$, where $Z = (1, z_{p-1}, \ldots, z_0)^T$. Furthermore, it is assumed that $d_\tau \le p$, without loss of generality.*

Assumption 4. *$F_t(\cdot)$ is absolutely continuous, $0 < f_t(u) < \infty$ on $\mathcal{U} = \{u : 0 < F_t(u) < 1\}$, and $f_t[F_t^{-1}(\tau)] > 0$.*

Assumptions 1–3 are regularity conditions used in Li et al. (2015), and Assumption 4 is necessary to derive the conditional quantile estimation (Koenker and Xiao (2006)).

Theorem 2. *Suppose that $E(|y_t|^{4+\varsigma}) < \infty$ for some $\varsigma > 0$, and matrix Ω_1 is positive-definite. If Assumptions 1–4 hold, then*

(a) $n(\widehat{r}_{L,\tau} - r_{L,\tau}^0) = O_p(1)$ *and* $n(\widehat{r}_{U,\tau} - r_{U,\tau}^0) = O_p(1)$*;*

(b) $\sqrt{n} \sup_{n(|r_{L,\tau} - r_{L,\tau}^0| + |r_{U,\tau} - r_{U,\tau}^0|) \le B} \|\widetilde{\theta}_{n,\tau}(r_{L,\tau}, r_{U,\tau}) - \widetilde{\theta}_{n,\tau}(r_{L,\tau}^0, r_{U,\tau}^0)\| = o_p(1)$ *for any fixed $0 < B < \infty$, where $\widetilde{\theta}_{n,\tau}(r_{L,\tau}, r_{U,\tau})$ is defined in the previous section;*

(c) $\sqrt{n}(\widehat{\theta}_{n,\tau} - \theta_\tau^0) \to_d N(0, \Sigma)$*, where* $\Sigma = \tau(1-\tau)\Omega_1^{-1}\Omega_0\Omega_1^{-1}$.

In real applications, we may be interested in the quantities of $\Xi\theta_\tau$, where Ξ is a known $k \times (2p+2)$ matrix with a full rank; see, for example, the generalized linear hypotheses in regression models. From Theorem 2, it holds that $\sqrt{n}\Xi(\widehat{\theta}_{n,\tau} - \theta_\tau^0) \to_d N(0, \Xi\Sigma\Xi^T)$; hence, we can design its inference tools accordingly.

The matrix Ω_1 in the asymptotic variance of $\widehat{\theta}_{n,\tau}$ involves the conditional density $f_t(\cdot)$. As in Koenker (2005) and Li, Li and Tsai (2015), we first consider a nonparametric method to estimate $f_t[F_t^{-1}(\tau)]$:

$$\widehat{f}_t[F_t^{-1}(\tau)] = \frac{2h}{\widehat{Q}_{y_t}(\tau + h|\mathcal{F}_{t-1}) - \widehat{Q}_{y_t}(\tau - h|\mathcal{F}_{t-1})},$$

where $\widehat{Q}_{y_t}(\tau|\mathcal{F}_{t-1}) = M_t(\widehat{\lambda}_{n,\tau})$. The matrices Ω_0 and Ω_1 can then be estimated by the sample averages, and hence the asymptotic variance Σ.

Alternatively, we may consider a bootstrap method to approximate the variance of $\widehat{\theta}_{n,\tau}$. From Theorems 1 and 2, without loss of generality, the parameters of $r_{L,\tau}$, $r_{U,\tau}$, and d_τ can be assumed to be known. By adopting the random weighting method of Rao and Zhao (1992) and Li, Li and Tsai (2015), we suggest the following bootstrapping procedure:

(a) Generate nonnegative independent and identically distributed (*i.i.d.*) random weights $\{\omega_t\}$ with both mean and variance equal one.

(b) Calculate

$$\widehat{\theta}^*_{n,\tau} = \operatorname{argmin} \sum_{t=1}^{n} \omega_t \rho_\tau[y_t - \widetilde{M}_t(\theta_\tau, \widehat{r}_{L,\tau}, \widehat{r}_{U,\tau}, \widehat{d}_\tau)],$$

where $(\widehat{r}_{L,\tau}, \widehat{r}_{U,\tau}, \widehat{d}_\tau)$ denotes the conditional quantile estimator of $(r_{L,\tau}, r_{U,\tau}, d_\tau)$.

(c) Repeat Steps (a) and (b) B times, and denote the resulting quantities by $\{\widehat{\theta}^{*(1)}_{n,\tau}, \ldots, \widehat{\theta}^{*(B)}_{n,\tau}\}$. The sample variance of $\{\widehat{\theta}^{*(k)}_{n,\tau} - \widehat{\theta}_{n,\tau}, 1 \leq k \leq B\}$ can then be used to approximate the variance of $\widehat{\theta}_{n,\tau}$.

Theorem 3. *Under the conditions of Theorem 2, it holds that, conditional on $y_1, \ldots, y_n$,*

$$\sqrt{n}(\widehat{\theta}^*_{n,\tau} - \widehat{\theta}_{n,\tau}) \to_d N(0, \Sigma)$$

in probability as $n \to \infty$.

Let $\widehat{p}_n = \operatorname{argmin}_{0 \leq p \leq p_{\max}} \mathrm{BIC}(p)$. Next, we give a theoretical justification for the BIC proposed in the previous section.

Theorem 4. *Under the conditions of Theorem 2, if $p_{\max} \geq p_0$, then $P\{\widehat{p}_n = p_0\} \to 1$ as $n \to \infty$, where p_0 is the true order; that is, $|\theta^0_{1,p_0}| + |\theta^0_{2,p_0}| \neq 0$.*

Using a method similar to that of the above theorem, we can show that minimizing the AIC tends to select an order that is greater than or equal to p_0.

4. Simulation Experiments

This section conducts three simulation experiments to evaluate the finite-sample performance of the conditional quantile estimation. In all experiments, we consider four quantiles, $\tau = 0.2$, 0.4, 0.6, and 0.8, and three sample sizes, $n = 100$, 200, and 500. There are 100 replications for each combination of quantile and sample size. The number of bootstrapped samples B is set to 1,000.

The data-generating process in the first experiment is

$$y_t = \begin{cases} \theta_{01}(U_t) + \theta_{11}(U_t) y_{t-1}, & R_t = 1, \\ \theta_{02}(U_t) + \theta_{12}(U_t) y_{t-1}, & R_t = 0, \end{cases} \tag{4.1}$$

with

$$R_t = \begin{cases} 1, & y_{t-2} \leq 1.12, \\ 0, & y_{t-2} > 1.85, \\ R_{t-1}, & \text{otherwise}, \end{cases}$$

where $\{U_t\}$ are *i.i.d.* standard uniform random variables over [0, 1], $\theta_{01}(x) = 0.85 + 0.15x$, $\theta_{11}(x) = 1/(e^{-x} + 1)$, $\theta_{02}(x) = 0.5$, and $\theta_{12}(x) = 1/(e^{-x} + e^{0.5})$. The conditional quantile estimation in Section 2 is applied with a (or b) being the 10th (or 90th) percentile of each sample. The asymptotic variances of $\widehat{\theta}_{n,\tau}$ are estimated using both the nonparametric method and a bootstrapping approximation. As in Koenker and Xiao (2006) and Li, Li and Tsai (2015), we choose the bandwidth of $3h_{HS}$ in the nonparametric method, where

$$h_{HS} = n^{-1/3} \left[\Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \right]^{2/3} \left[\frac{1.5\phi^2 \Phi^{-1}(\tau)}{2\left[\Phi^{-1}(\tau)\right]^2 + 1} \right]^{1/3},$$

and $\Phi(\cdot)$ is the distribution of the standard normal distribution; see Hall and Sheather (1988). Tables 1 and 2 list the estimation results for $\tau = \{0.2, 0.4\}$ and $\{0.6, 0.8\}$, respectively. They include the bias (BIAS), and the asymptotic variances calculated using the nonparametric method (ASD) and the bootstrap method (BSD). The results show that both the bias and the ESDs get smaller as the sample size increases. Despite that, in most cases, the BSDs are slightly above the ASDs. In addition, they are close to each other and close to the ESDs when the sample size is as small as $n = 200$.

The second experiment employs the same data-generating process as that in (4.1), with the regime indicator function

Table 1. Estimation results for the HAR models in the first experiment with $\tau = 0.2$ and 0.4.

n		$\widehat{\theta}_{01,\tau}$	$\widehat{\theta}_{11,\tau}$	$\widehat{\theta}_{02,\tau}$	$\widehat{\theta}_{12,\tau}$	$\widehat{r}_{L,\tau}$	$\widehat{r}_{U,\tau}$
				$\tau = 0.2$			
100	BIAS	0.0035	0.0029	0.0012	−0.0002	−0.0085	−0.0204
	ESD	0.0965	0.0693	0.0358	0.0269	0.0126	0.0265
	ASD	0.0970	0.0680	0.0392	0.0295		
	BSD	0.1083	0.0744	0.0413	0.0307		
200	BIAS	0.0095	−0.0031	0.0012	−0.0008	−0.0051	−0.0068
	ESD	0.0652	0.0471	0.0285	0.0216	0.0080	0.0146
	ASD	0.0687	0.0481	0.0278	0.0209		
	BSD	0.0731	0.0512	0.0291	0.0216		
500	BIAS	−0.0006	0.0010	0.0024	−0.0009	−0.0062	−0.0060
	ESD	0.0476	0.0317	0.0165	0.0124	0.0044	0.0116
	ASD	0.0436	0.0305	0.0176	0.0133		
	BSD	0.0474	0.0328	0.0176	0.0132		
				$\tau = 0.4$			
100	BIAS	0.0033	−0.0016	−0.0015	0.0001	−0.0071	−0.0136
	ESD	0.1271	0.0878	0.0382	0.0290	0.0125	0.0247
	ASD	0.1167	0.0814	0.0445	0.0334		
	BSD	0.1192	0.0828	0.0444	0.0329		
200	BIAS	−0.0072	0.0050	−0.0031	0.0013	−0.0046	−0.0079
	ESD	0.0854	0.0590	0.0322	0.0244	0.0094	0.0185
	ASD	0.0828	0.0578	0.0317	0.0239		
	BSD	0.0888	0.0615	0.0327	0.0247		
500	BIAS	−0.0117	0.0084	0.0055	−0.0035	−0.0026	−0.0007
	ESD	0.0512	0.0360	0.0202	0.0153	0.0063	0.0100
	ASD	0.0524	0.0367	0.0200	0.0151		
	BSD	0.0533	0.0369	0.0207	0.0156		

$$R_t = \begin{cases} 1, & y_{t-2} \le 2.55, \\ 0, & y_{t-2} > 3.57, \\ R_{t-1}, & \text{otherwise}, \end{cases}$$

where $\theta_{01}(x) = F^{-1}_{\chi^2_1}[1/(e^{-x}+1)]$, $\theta_{11}(x) = 1$, $\theta_{02}(x) = 0.5$, $\theta_{12}(x) = 1/(e^{-x}+1)$, and $F^{-1}_{\chi^2_1}(\cdot)$ is the quantile function of the χ^2_1 distribution with one degree of freedom. The generated sequences are all nonnegative. There is a unit root in the structure at the lower regime, but the overall model remains stationary (Koenker and Xiao (2004)). All other settings are the same as those in the first experiment. The estimation results are presented in Tables 3 and 4 for $\tau = \{0.2, 0.4\}$ and $\{0.6, 0.8\}$, respectively. Similar findings are observed.

Table 2. Estimation results for the HAR models in the first experiment with $\tau = 0.6$ and 0.8.

n		$\widehat{\theta}_{01,\tau}$	$\widehat{\theta}_{11,\tau}$	$\widehat{\theta}_{02,\tau}$	$\widehat{\theta}_{12,\tau}$	$\widehat{r}_{L,\tau}$	$\widehat{r}_{U,\tau}$
				$\tau = 0.6$			
100	BIAS	−0.0026	−0.0033	−0.0009	0.0005	−0.0079	−0.0142
	ESD	0.1105	0.0753	0.0375	0.0266	0.0133	0.0259
	ASD	0.1132	0.0792	0.0406	0.0306		
	BSD	0.1141	0.0786	0.0424	0.0313		
200	BIAS	−0.0047	0.0016	0.0042	−0.0026	−0.0008	−0.0049
	ESD	0.0845	0.0577	0.0302	0.0225	0.0081	0.0170
	ASD	0.0800	0.0558	0.0288	0.0217		
	BSD	0.0822	0.0567	0.0302	0.0227		
500	BIAS	−0.0001	−0.0005	−0.0005	0.0005	0.0021	0.0031
	ESD	0.0539	0.0385	0.0202	0.0153	0.0059	0.0057
	ASD	0.0508	0.0355	0.0182	0.0137		
	BSD	0.0524	0.0363	0.0185	0.0138		
				$\tau = 0.8$			
100	BIAS	−0.0039	−0.0015	0.0095	−0.0088	−0.0049	−0.0134
	ESD	0.0923	0.0637	0.0320	0.0255	0.0126	0.0264
	ASD	0.0885	0.0616	0.0302	0.0227		
	BSD	0.0943	0.0661	0.0330	0.0253		
200	BIAS	−0.0007	−0.0014	0.0004	−0.0010	0.0006	−0.0044
	ESD	0.0667	0.0440	0.0206	0.0151	0.0082	0.0165
	ASD	0.0627	0.0437	0.0211	0.0158		
	BSD	0.0635	0.0445	0.0224	0.0168		
500	BIAS	−0.0042	0.0013	0.0011	−0.0010	0.0042	0.0047
	ESD	0.0402	0.0272	0.0124	0.0094	0.0043	0.0040
	ASD	0.0397	0.0277	0.0134	0.0101		
	BSD	0.0416	0.0290	0.0141	0.0107		

In the third experiment, we perform a conditional least squares estimation (Li et al. (2015)) on the samples generated in the first two experiments. The estimation results are given in Table 5, and both the bias and the ESDs decrease as the sample size increases. Note that the data-generating process in (4.1) can be rewritten in a mean regression form,

$$y_t = \begin{cases} \theta_{01} + \theta_{11} y_{t-1} + \varepsilon_{1t}, & R_t = 1, \\ \theta_{02} + \theta_{12} y_{t-1} + \varepsilon_{2t}, & R_t = 0, \end{cases}$$

where $\theta_{ij} = E[\theta_{ij}(U_t)]$ for $0 \leq i \leq 1$ and $1 \leq j \leq 2$, $\varepsilon_{jt} = \theta_{0j}(U_t) - \theta_{0j} + [\theta_{1j}(U_t) - \theta_{1j}]y_{t-1}$ for $1 \leq j \leq 2$, and $\{(\varepsilon_{1t}, \varepsilon_{2t}), \mathcal{F}_t\}$ is a martingale difference sequence. As a result, both the consistency and the asymptotic normality can be obtained.

Table 3. Estimation results for the HAR models in the second experiment with $\tau = 0.2$ and 0.4.

n		$\widehat{\theta}_{01,\tau}$	$\widehat{\theta}_{11,\tau}$	$\widehat{\theta}_{02,\tau}$	$\widehat{\theta}_{12,\tau}$	$\widehat{r}_{L,\tau}$	$\widehat{r}_{U,\tau}$
				$\tau = 0.2$			
100	BIAS	0.0225	−0.0039	0.0209	−0.0045	−0.0349	−0.0457
	ESD	0.1294	0.0424	0.1441	0.0460	0.0397	0.0582
	ASD	0.1298	0.0425	0.1518	0.0509		
	BSD	0.1425	0.0466	0.1595	0.0524		
200	BIAS	0.0034	−0.0001	0.0035	−0.0023	−0.0139	−0.0183
	ESD	0.0918	0.0298	0.0993	0.0334	0.0228	0.0313
	ASD	0.0918	0.0301	0.1080	0.0361		
	BSD	0.0948	0.0319	0.1118	0.0370		
500	BIAS	−0.0080	0.0037	−0.0080	0.0041	−0.0058	−0.0023
	ESD	0.0547	0.0186	0.0692	0.0234	0.0142	0.0097
	ASD	0.0576	0.0189	0.0678	0.0227		
	BSD	0.0608	0.0203	0.0702	0.0236		
				$\tau = 0.4$			
100	BIAS	0.0281	−0.0059	−0.0249	0.0070	−0.0273	−0.0385
	ESD	0.1954	0.0648	0.1564	0.0552	0.0392	0.0576
	ASD	0.1841	0.0605	0.1757	0.0589		
	BSD	0.1965	0.0640	0.1853	0.0610		
200	BIAS	−0.0085	0.0029	0.0005	0.0001	−0.0106	−0.0210
	ESD	0.1384	0.0446	0.1286	0.0445	0.0216	0.0336
	ASD	0.1292	0.0424	0.1257	0.0421		
	BSD	0.1339	0.0442	0.1313	0.0440		
500	BIAS	−0.0049	0.0017	0.0022	−0.0003	−0.0019	−0.0019
	ESD	0.0752	0.0257	0.0849	0.0274	0.0131	0.0104
	ASD	0.0816	0.0268	0.0804	0.0269		
	BSD	0.0841	0.0276	0.0821	0.0274		

5. An Empirical Example

We study unemployment rates, because these have important implications for economic policymaking. Many researchers have studied the asymmetric dynamics in the response of unemployment to economic expansions and contractions. Koenker and Xiao (2006) used the quantile AR model to analyze US quarterly and annual unemployment rates, finding that the estimated AR roots vary over quantiles. Galvao, Montes-Rojas and Olmo (2011) carried out a thorough study of the asymmetric dynamics of the conditional distribution of US monthly unemployment growth after World War II. The study is based on the threshold AR model, and a stronger asymmetric persistence is suggested in higher quantiles. At the same time, hysteresis has been studied extensively in the literature, and

Table 4. Estimation results for the HAR models in the second experiment with $\tau = 0.6$ and 0.8.

n		$\widehat{\theta}_{01,\tau}$	$\widehat{\theta}_{11,\tau}$	$\widehat{\theta}_{02,\tau}$	$\widehat{\theta}_{12,\tau}$	$\widehat{r}_{L,\tau}$	$\widehat{r}_{U,\tau}$
				$\tau = 0.6$			
100	BIAS	−0.0010	0.0002	−0.0017	−0.0017	−0.0316	−0.0383
	ESD	0.2095	0.0658	0.1680	0.0549	0.0436	0.0487
	ASD	0.2061	0.0678	0.1688	0.0566		
	BSD	0.2097	0.0688	0.1784	0.0606		
200	BIAS	−0.0292	0.0087	−0.0336	0.0096	−0.0071	−0.0093
	ESD	0.1373	0.0456	0.1214	0.0429	0.0216	0.0278
	ASD	0.1456	0.0477	0.1199	0.0402		
	BSD	0.1439	0.0473	0.1293	0.0428		
500	BIAS	0.0099	−0.0041	0.0025	−0.0005	0.0014	−0.0009
	ESD	0.0997	0.0344	0.0796	0.0276	0.0104	0.0074
	ASD	0.0920	0.0302	0.0766	0.0257		
	BSD	0.0958	0.0316	0.0798	0.0262		
				$\tau = 0.8$			
100	BIAS	−0.0042	0.0006	−0.0121	0.0019	−0.0204	−0.0384
	ESD	0.2004	0.0684	0.1438	0.0458	0.0357	0.0549
	ASD	0.1881	0.0615	0.1323	0.0443		
	BSD	0.1947	0.0639	0.1375	0.0454		
200	BIAS	0.0023	−0.0027	0.0023	−0.0020	−0.0084	−0.0153
	ESD	0.1436	0.0452	0.1057	0.0357	0.0188	0.0337
	ASD	0.1328	0.0435	0.0928	0.0311		
	BSD	0.1349	0.0440	0.1006	0.0335		
500	BIAS	−0.0077	0.0004	−0.0117	0.0037	0.0036	−0.0016
	ESD	0.0838	0.0274	0.0631	0.0209	0.0079	0.0087
	ASD	0.0841	0.0276	0.0589	0.0198		
	BSD	0.0877	0.0288	0.0617	0.0205		

has been confirmed for the unemployment rate (Blanchard and Summers (1986); Jaeger and Parkinson (1994); Perez-Alonso and Sanzo (2011)). As a result, the HAR model may be more suitable for modeling such asymmetry.

This section considers the growth rate of US monthly unemployment rates, rather than the unemployment rates themselves, as the mean reverting behavior (Galvao, Montes-Rojas and Olmo (2011)). The study period is January 1948 to December 2007; the time plot is presented in Figure 1. We consider eight different quantiles $\tau = \{0.05, 0.10, 0.25, 0.40, 0.60, 0.75, 0.90, 0.95\}$. The BIC at (2.4) is employed to select the values of p and d, with $p_{\max} = d_{\max} = 5$; the selection results are given in Table 6. It can be seen that $p = 1$ is selected for all τ except 0.05 and 0.90, whereas $d = 1$ is chosen in all except three quantiles. For

Table 5. Estimation results for the HAR models under a mean regression in the first and second experiments.

n		$\widehat{\theta}_{01,\tau}$	$\widehat{\theta}_{11,\tau}$	$\widehat{\theta}_{02,\tau}$	$\widehat{\theta}_{12,\tau}$	$\widehat{r}_{L,\tau}$	$\widehat{r}_{U,\tau}$
				First Experiment			
100	BIAS	−0.0078	0.0059	−0.0045	0.0021	−0.0088	−0.0189
	ESD	0.0733	0.0509	0.0300	0.0230	0.0106	0.0238
200	BIAS	−0.0032	0.0034	−0.0019	0.0017	−0.0036	−0.0045
	ESD	0.0516	0.0387	0.0192	0.0137	0.0068	0.0084
500	BIAS	0.0039	−0.0035	−0.0023	0.0023	−0.0004	−0.0020
	ESD	0.0334	0.0229	0.0149	0.0116	0.0018	0.0053
				Second Experiment			
100	BIAS	−0.0268	0.0079	−0.0081	0.0016	−0.0382	−0.0544
	ESD	0.1220	0.0408	0.1236	0.0392	0.0453	0.0560
200	BIAS	0.0012	0.0003	−0.0015	0.0012	−0.0153	−0.0182
	ESD	0.0848	0.0278	0.0848	0.0283	0.0202	0.0251
500	BIAS	0.0076	−0.0025	0.0037	−0.0009	−0.0017	−0.0048
	ESD	0.0586	0.0185	0.0527	0.0178	0.0095	0.0094

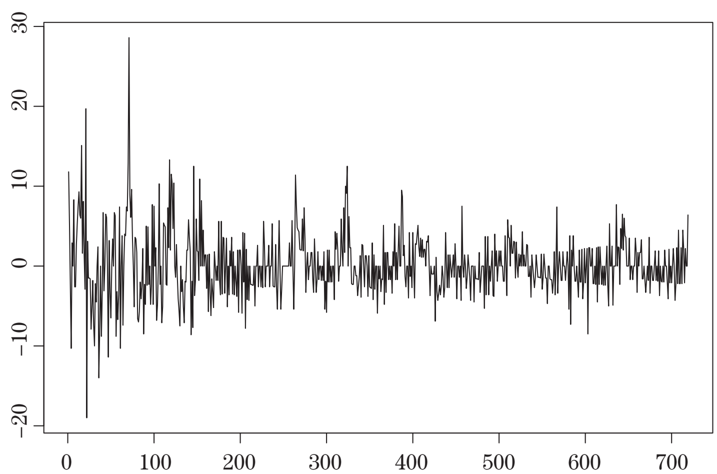

Figure 1. Growth rates, in percentages, of US monthly unemployment rates from January 1948 to December 2007.

ease of comparison, we fix $p = d = 1$. The estimation results are listed in Table 7 and Figure 2.

The fitted intercept for the upper regime crosses zero at $\tau = 0.40$, but crosses zero at $\tau = 0.60$ for the lower regime; both are monotonically increasing with τ. The slopes of the lower regime are all significantly greater than zero, which suggests strong serially correlated behavior of the unemployment rates for this

Table 6. Values of p and d selected by the BIC.

τ	0.05	0.1	0.25	0.4	0.6	0.75	0.9	0.95
p	4	1	1	1	1	1	2	1
d	2	1	1	2	1	1	1	2

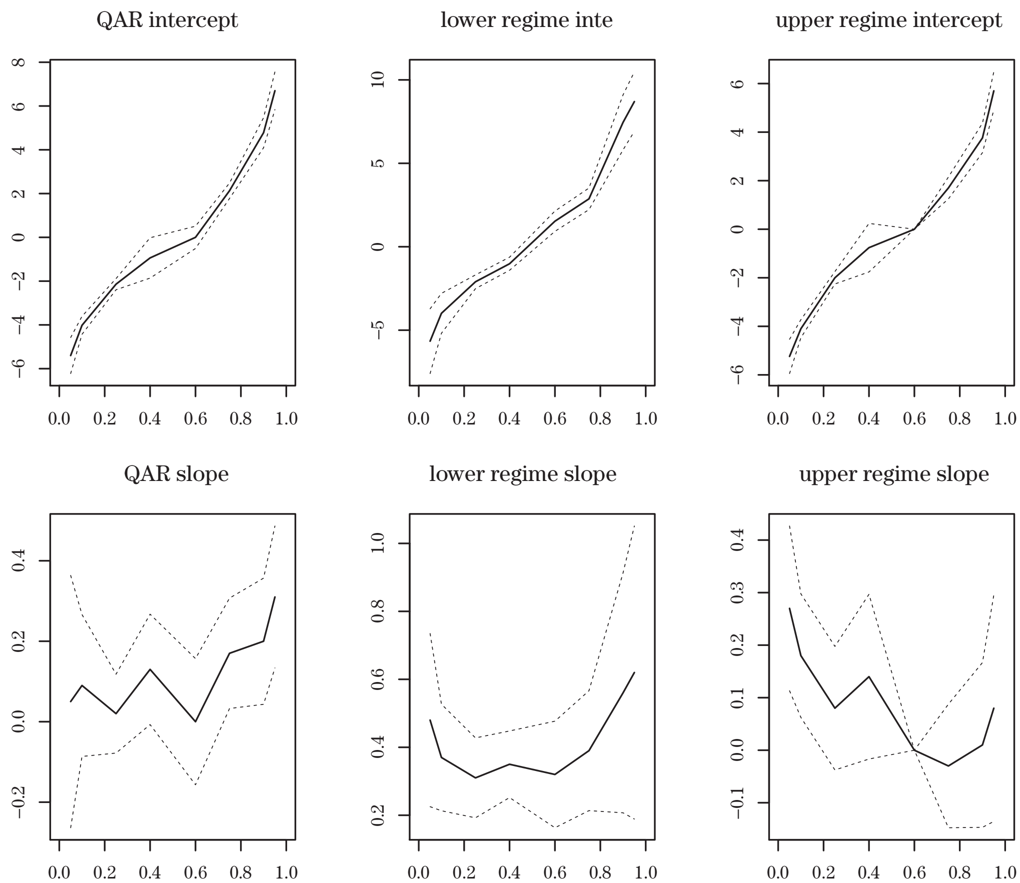

Figure 2. Estimated coefficients (solid lines) and their 95% confidence intervals (dotted lines) for the AR (left panel) and HAR (middle and right panels) models from the conditional quantile estimation.

regime. However, the slopes of the upper regime are all nonsignificant except for the lower quantiles $\tau = 0.05$ and 0.10; thus, the return series of the unemployment rate exhibits memoryless behavior for this regime. This observation is in line with the economic intuition that economic growth is usually considered the "normal" state of an economy, whereas economic recessions are considered anomalies, and should not last long. Therefore, for most of the sample period, the return series

Table 7. Estimated coefficients, together with standard errors in parentheses, of the HAR model from the conditional quantile estimation, and the percentages of observations (PO) falling within the hysteresis zone.

τ	$\widehat{\theta}_{01,\tau}$	$\widehat{\theta}_{11,\tau}$	$\widehat{\theta}_{02,\tau}$	$\widehat{\theta}_{12,\tau}$	$\widehat{r}_{L,\tau}$	$\widehat{r}_{U,\tau}$	PO
0.05	−5.66	0.48	−5.24	0.27	−1.30	3.65	0.47
	(0.99)	(0.13)	(0.36)	(0.08)			
0.1	−3.99	0.37	−4.10	0.18	−1.30	3.65	0.47
	(0.61)	(0.08)	(0.19)	(0.06)			
0.25	−2.10	0.31	−2.00	0.08	−1.70	0.00	0.06
	(0.21)	(0.06)	(0.13)	(0.06)			
0.4	−1.02	0.35	−0.76	0.14	−1.70	0.00	0.06
	(0.20)	(0.05)	(0.51)	(0.08)			
0.6	1.53	0.32	0.00	0.00	−1.30	3.10	0.44
	(0.31)	(0.08)	(0.00)	(0.00)			
0.75	2.87	0.39	1.70	−0.03	−1.30	2.30	0.38
	(0.33)	(0.09)	(0.23)	(0.06)			
0.9	7.45	0.56	3.76	0.01	1.30	3.65	0.21
	(0.84)	(0.18)	(0.31)	(0.08)			
0.95	8.70	0.62	5.70	0.08	1.30	3.65	0.21
	(0.91)	(0.22)	(0.40)	(0.11)			

Note: "0.00" refers to a value smaller than 0.005.

Table 8. Values of the BIC for the fitted HAR and TAR models with two (TAR2) and three regimes (TAR3).

τ	0.05	0.1	0.25	0.4	0.6	0.75	0.9	0.95
HAR	−694	−344	57	204	198	114	−243	−584
TAR2	−662	−322	65	213	221	127	−224	−565
TAR3	−685	−333	65	210	221	128	−232	−577

of the unemployment rate tend to exhibit strong serially correlated behavior at relatively low levels (lower regime). Occasional large shocks might push the series into the upper regime, but they will exit quickly owing to the lack of memory. However, with the presence of the hysteresis zone, the series will not immediately fall back into the lower regime, but instead will encounter delays, or may even switch back and forth, leading to a period of high unemployment that lasts longer than expected. Hence, by explicitly incorporating a hysteresis zone, our model leads to an interpretation that dramatically differs from that of the quantile threshold autoregressive (TAR) model of Galvao, Montes-Rojas and Olmo (2011), and is more consistent with the economic intuition.

We also compare the HAR and TAR models with two and three regimes in terms of the BIC; the results are given in Table 8. The evidence of hysteresis is further reinforced by the observation that the BIC of the HAR model is the lowest at all quantiles. We conclude that the HAR model is more suitable than the TAR model for interpreting unemployment rates. Note that the BIC for the HAR is supposed to be smaller than that for the TAR model with two regimes, because the former includes the latter as a special case.

For the sake of comparison, we also fit the quantile AR model in Koenker and Xiao (2006) to the data. The order is chosen as one, in line with the choice of p in the HAR models; the fitted coefficients are presented in Figure 2. It can be seen that the slope parameters of the fitted model do not significantly differ from zero for lower quantiles, but are significantly positive for the upper quantiles, $\tau = 0.75, 0.90$, and 0.95. This implies the presence of asymmetric dynamics, and hence the necessity of a regime-switching model; see also Figure 4 in Koenker and Xiao (2006). Moreover, we calculate the conditional least squares estimation for the AR and HAR models,

$$y_t = 0.14_{0.17} \ + \ 0.13_{0.04} y_{t-1} + \varepsilon_t,$$

and

$$y_t = \begin{cases} 0.15_{0.19} \ + \ 0.44_{0.07} y_{t-1} + \varepsilon_{1t}, & R_t = 1, \\ -0.02_{0.19} \ + \ 0.05_{0.06} y_{t-1} + \varepsilon_{2t}, & R_t = 0, \end{cases}$$

respectively, with $d = 2$, $r_L = -1.80$, and $r_U = 0.00$, where the standard errors are given in the subscripts. For the fitted HAR model, the slope parameter of the lower regime is significant and positive, whereas that of the upper regime, similarly to its quantile counterparts, is not significantly different from zero. In fact, they can be considered the averaged values over all quantiles; see also the third experiment in the previous section. Finally, the fitted AR model seems to offer a compromise between these two structures in the HAR model.

6. Conclusion

This study develops a conditional quantile estimation for HAR models, which is useful for modeling economic time series with hysteresis, for example, unemployment rates. This estimation gives us greater flexibility in understanding the hysteresis patterns at different quantiles. The asymptotic behaviors of the estimators are established.

Several open problems related to the conditional quantile estimation for HAR models remain, which we leave to future research. First, an important task is to test for the existence of the threshold. Galvao et al. (2014) and Zhang, Wang and Zhu (2014) proposed tests for the threshold effect at some quantile levels, and Zhu, Yu and Li (2014) conducted a quasi-likelihood ratio test for the linearity against hysteresis AR processes. It should be feasible to construct a test for $\theta^0_{1,\tau} \neq \theta^0_{2,\tau}$ in Assumption 1 by following Kato (2009), Galvao et al. (2014), Zhang, Wang and Zhu (2014), and Zhu, Yu and Li (2014).

Second, we provide theoretical justifications for the super-consistency of the estimated boundary parameters, $\widehat{r}_{L,\tau}$ and $\widehat{r}_{U,\tau}$, only; thus, it is of interest to derive their asymptotic distributions, as in Li et al. (2015) and Kuan, Michalopoulos and Xiao (2017). Third, we would like to extend our theoretical results from a fixed τ to a close set $\mathcal{I} \in (0,1)$; the theoretical tools in Kato (2009) may be of assistance here. Finally, it is important to construct a diagnostic tool to check the adequacy of the fitted HAR model using the conditional quantile estimation.

Supplementary Material

The online Supplementary Material contains the proofs of Theorems 1 to 4.

Acknowledgments

We thank the editor, an associate editor, and two anonymous referees for their valuable comments and suggestions. This research was partially supported by Hong Kong RGC grants 17304417 and 17304617, NSFC grant 11601313, and National Statistical Scientific Research Project 2017LY32.

References

Amable, B., Lordon, J., Henry, F. and Topol, R. (1995). Hysteresis revisited: a methodological approach. In *The Natural Rate of Unemployment*, 153–180. Cambridge University Press.

Blanchard, O. J. and Summers, L. H. (1986). Hysteresis and the european unemployment problem. In *NBER Macroeconomics Annual 1986* **1**, 15–90. MIT Press.

Brunello, G. (1990). Hysteresis and "the Japanese unemployment problem" : a preliminary investigation. *Oxford Economic Papers* **42**, 483–500.

Cai, Y. (2010). Forecasting for quantile self-exciting threshold autoregressive time series models. *Biometrika* **97**, 199–208.

Cai, Y. and Stander, J. (2008). Quantile self-exciting threshold autoregressive time series models. *Journal of Time Series Analysis* **29**, 186–202.

Chan, K. S. and Tong, H. (1986). On estimating thresholds in autoregressive models. *The Annals of Statistics* **7**, 179–190.

Galvao, A. F., Kato, K., Montes-Rojas, G. and Olmo, J. (2014). Testing linearity against threshold effects: uniform inference in quantile regression. *Annals of the Institute of Statistical*

Mathematics **66**, 413–439.

Galvao, A. F., Montes-Rojas, G. and Olmo, J. (2011). Threshold quantile autoregressive models. *Journal of Time Series Analysis* **32**, 253–267.

Hall, P. and Sheather, S. J. (1988). On the distribution of a studentized quantile. *Journal of Royal Statistical Society, Series B (Statistical Methodology)* **50**, 381–391.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**, 357–384.

Jaeger, A. and Parkinson, M. (1994). Some evidence on hysteresis in unemployment rates. *European Economic Review* **38**, 329–342.

Kato, K. (2009). Asymptotics for argmin processes: convexity arguments. *Journal of Multivariate Analysis* **100**, 1816–1829.

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

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica* **46**, 33–49.

Koenker, R. and Xiao, Z. (2004). Unit root quantile autoregression inference. *Journal of the American Statistical Association* **99**, 775–787.

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

Kuan, C.-M., Michalopoulos, C. and Xiao, Z. (2017). Quantile regression on quantile ranges - a threshold approach. *Journal of Time Series Analysis* **38**, 99–119.

Lee, E. R., Noh, H. and Park, B. U. (2014). Model selection via bayesian information criterion for quantile regression models. *Journal of the American Statistical Association* **109**, 216–229.

Li, G., Guan, B., Li, W. K. and Yu, P. L. H. (2015). Hysteretic autoregressive time series models. *Biometrika* **102**, 717–723.

Li, G. and Li, W. K. (2008). Testing for threshold moving average with conditional heteroscedasticity. *Statistica Sinica* **18**, 647–665.

Li, G. and Li, W. K. (2011). Testing a linear time series model against its threshold extension. *Biometrika* **98**, 243–250.

Li, G., Li, Y. and Tsai, C.-L. (2015). Quantile correlations and quantile autoregressive modeling. *Journal of the American Statistical Association* **110**, 246–261.

McCulloch, R. E. and Tsay, R. S. (1994). Statistical inference of macroeconomic time series via markov switching models. *Journal of Time Series Analysis* **15**, 523–539.

Perez-Alonso, A. and Sanzo, S. D. (2011). Unemployment and hysteresis: a nonlinear unobserved components approach. *Studies in Nonlinear Dynamics and Econometrics* **15**, 1–27.

Rao, C. R. and Zhao, L. C. (1992). Approximation to the distribution of m-estimates in linear models by randomly weighted bootstrap. *Sankhya* **54**, 323–331.

Roed, B. K. (2002). Unemployment hysteresis and the natural rate of vacancies. *Empirical Economics* **27**, 687–704.

Song, F. M. and Wu, Y. (1997). Hysteresis in unemployment: evidence from 48 states. *Economic Inquiry* **35**, 235–244.

Song, F. M. and Wu, Y. (1998). Hysteresis unemployment: evidence from oecd countries. *The Quarterly Review of Economics and Finance* **38**, 181–192.

Tong, H. (1990). *Non-Linear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford, UK.

Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal* *of the Royal Statistical Society, Series B (Statistical Methodology)* **42**, 245–292.

Wu, S. and Chen, R. (2007). Threshold variable determination and threshold variable driven switching autoregressive models. *Statistica Sinica* **17**, 241–264.

Zhang, L., Wang, H. J. and Zhu, Z. (2014). Testing for change points due to a covariate threshold in regression quantiles. *Statistica Sinica* **24**, 1859–1877.

Zhu, K., Yu, P. L. H. and Li, W. K. (2014). Testing for the buffered autoregressive processes. *Statistica Sinica* **24**, 971–984.

College of Mathematics and Information Engineering, Jiaxing University,Jiaxing, China.

E-mail: dgli@mail.zjxu.edu.cn

Department of Statistics and Actuarial Science, Univesity of Hong Kong, Hong Kong.

E-mail: xkd17@connect.hku.hk

School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, China.

E-mail: zhang.liwen@mail.shufe.edu.cn

Department of Statistics and Actuarial Science, Univesity of Hong Kong, Hong Kong.

E-mail: hrntlwk@hku.hk

Department of Statistics and Actuarial Science, Univesity of Hong Kong, Hong Kong.

E-mail: gdli@hku.hk

(Received July 2017; accepted June 2018)