Statistica Sinica **26** (2016), 1555-1567
doi:http://dx.doi.org/10.5705/ss.2014.098t

ON BUFFERED THRESHOLD GARCH MODELS

Pak Hang Lo, Wai Keung Li, Philip L. H. Yu and Guodong Li

University of Hong Kong

Abstract: This paper proposes a conditional heteroscedastic model with a new piecewise linear structure such that the regime-switching mechanism has a buffer zone where regime-switching is delayed. Gaussian quasi-maximum likelihood estimation (QMLE) is considered, and its asymptotic behaviors, including strong consistency and the asymptotic distribution, are derived. Its finite sample performance is evaluated by Monte Carlo simulation experiments, and an empirical example is reported to give further support to the new model.

Key words and phrases: Buffered threshold model, GARCH model, QMLE, threshold model.

1. Introduction

The threshold time series model has become a very successful type of nonlinear model since its introduction by Tong (1978) and Tong and Lim (1980). See also Chan (1993), Tsay (1998), Ling and Tong (2005), etc. Meanwhile, the generalized autoregressive conditional heteroscedastic (GARCH) model (Engle (1982); Bollerslev (1986)) is another classical nonlinear time series model in interpreting the clustering phenomenon in volatilities of many asset prices. It was then natural for Li and Li (1996) to introduce the threshold autoregressive conditional heteroscedastic (ARCH) model by combining these two useful ideas. Liu, Li, and Li (1997) further extended the threshold ARCH model to a more general threshold GARCH model. These two models have been shown to be very useful in modeling and interpreting the asymmetry in volatilities of asset prices (Brooks (2001)).

While threshold models have achieved a huge success, they have also been observed to have bad performance near the boundaries between different regimes (Wu and Chen (2007)). The smooth-transition threshold autoregressive (AR) model (Chan and Tong (1986); van Dijk, Terasvirta, and Franses (2002)) can reduce this problem to some extent, but it usually requires more observations in estimating the transition function, and may not perform well for the kind of changes resembling quantum jumps. Moreover, the model does not have the simple structure of a piecewise linear specification. The discrete-state Markov switching AR model (Hamilton (1989); McCulloch and Tsay (1994)) does not encounter this problem since its regime switching is completely controlled by a latent random variable. However, it may not be easy to find a physical interpretation of the fitted model although it enjoys certain flexibility in switching regimes. Wu and Chen (2007) considered a threshold AR model with the switching mechanism jointly driven by observable variables and a latent variable; however, it still lacks a physical interpretation due to the unobservable latent variable.

For a threshold model, say with two regimes, there is a single threshold where the model switches its probabilistic structure. However, this may not be the case in reality. As an illustrative example, consider the asymmetry in volatilities of asset prices: empirical evidence has shown that asset prices have different volatility structures for good news and bad news (Bekaert and Wu (2000)). When the return of an asset price up-crosses a certain positive threshold r_U, the market can assert the coming of a good news. While bad news is not confirmed until the return down-crosses another negative threshold r_L. The interval $(r_L, r_U]$ acts as a buffer zone. There is no news coming when the return stays in the buffer zone, and the volatility structure is also supposed to keep unchanged. To capture this type of pattern, Li et al. (2015) discussed a threshold autoregressive model and a new regime-switching mechanism. Specifically, the time series is at the "lower" regime when threshold variable $z_t \le r_L$; at the "upper" regime when $z_t > r_U$, and it keeps the regime unchanged as long as z_t falls in $(r_L, r_U]$.

We adopt the idea in Li et al. (2015) to propose a threshold conditional heteroscedastic model with a more flexible regime-switching mechanism. The model is introduced in Section 2; it is called the buffered threshold GARCH model for simplicity. We derive the asymptotic behavior of its Gaussian quasi-maximum likelihood estimation (QMLE), including strong consistency and the asymptotic distribution, in Section 3. Section 4 considers several Monte Carlo simulation experiments to evaluate the finite sample performance of the Gaussian QMLE, and an empirical example is reported in Section 5 to give further support to the new model. Section 6 gives a short conclusion.

2. Buffered Threshold GARCH Models

Consider a two-regime buffered threshold GARCH model,

$$y_t = \varepsilon_t \sigma_t, \qquad \sigma_t^2 = \begin{cases} \omega^{(1)} + \sum_{i=1}^{q} \alpha_i^{(1)} y_{t-i}^2 + \sum_{j=1}^{p} \beta_j^{(1)} \sigma_{t-j}^2, & \text{if } R_t = 1, \\ \omega^{(2)} + \sum_{i=1}^{q} \alpha_i^{(2)} y_{t-i}^2 + \sum_{j=1}^{p} \beta_j^{(2)} \sigma_{t-j}^2, & \text{if } R_t = 0, \end{cases} \tag{2.1}$$

with the regime indicator

$$R_t = \begin{cases} 1, & \text{if } y_{t-d} \le r_L, \\ R_{t-1}, & \text{if } r_L < y_{t-d} \le r_U, \\ 0, & \text{if } y_{t-d} > r_U, \end{cases} \tag{2.2}$$

where $\omega^{(l)} > 0$, $\alpha_i^{(l)} \ge 0$ and $\beta_j^{(l)} \ge 0$ with $l = 1$ and 2, $\{\varepsilon_t\}$ are independently and identically distributed (*i.i.d.*) random variables with mean zero and variance one, r_L and r_U are the boundary (or threshold) parameters of the buffer zone $(r_L, r_U]$, and the positive integer d is the delay parameter. When $r_L = r_U$, models (2.1) and (2.2) reduce to the traditional threshold GARCH model (Liu, Li, and Li (1997)).

It is implied by (2.2) that

$$\begin{aligned} R_t &= I(y_{t-d} \le r_L) + I(r_L < y_{t-d} \le r_U) R_{t-1} \\ &= I(y_{t-d} \le r_L) + \sum_{j=0}^{\infty} \prod_{i=0}^{j} I(r_L < y_{t-d-i} \le r_U) I(y_{t-d-j-1} \le r_L) \end{aligned}$$

in the almost sure sense; see also Li et al. (2015). When $r_L < r_U$, the regime indicator R_t depends on past observations infinitely far away. This makes the buffered threshold model different from the traditional ones (Tong (1990); Hansen (2000)).

For the special case of (2.1) and (2.2) with $p = q = d = 1$, we have that

$$\begin{aligned} R_t &= I(\sigma_{t-1}\varepsilon_{t-1} \le r_L) + I(r_L < \sigma_{t-1}\varepsilon_{t-1} \le r_U) R_{t-1} := g_2(\sigma_{t-1}^2, R_{t-1}, \varepsilon_{t-1}), \\ \sigma_t^2 &= (\omega^{(1)} + \alpha^{(1)}\sigma_{t-1}^2\varepsilon_{t-1}^2 + \beta^{(1)}\sigma_{t-1}^2) R_t + (\omega^{(2)} + \alpha^{(2)}\sigma_{t-1}^2\varepsilon_{t-1}^2 + \beta^{(2)}\sigma_{t-1}^2)(1 - R_t) \\ &:= g_1(\sigma_{t-1}^2, R_{t-1}, \varepsilon_{t-1}), \end{aligned}$$

where the subscripts of $\alpha_1^{(i)}$ and $\beta_1^{(i)}$ are suppressed without confusion; see Francq and Zakoïan (2006) and Meitz and Saikkonen (2008). If $\boldsymbol{\sigma}_t^2 = (\sigma_t^2, R_t)'$, it holds that $\boldsymbol{\sigma}_t^2 = \mathbf{g}(\boldsymbol{\sigma}_{t-1}^2, \varepsilon_{t-1}) = (g_1(\sigma_{t-1}^2, R_{t-1}, \varepsilon_{t-1}), g_2(\sigma_{t-1}^2, R_{t-1}, \varepsilon_{t-1}))'$. As a result, $\{\boldsymbol{\sigma}_t^2\}$ is a homogenous Markov chain.

Theorem 1. *Suppose ε_t has a density function positive everywhere on $\mathbb{R}$. If $\max\{\alpha^{(1)}, \alpha^{(2)}\} + \max\{\beta^{(1)}, \beta^{(2)}\} < 1$, then $\{\boldsymbol{\sigma}_t^2\}$ is geometrically ergodic, hence the geometric ergodicity of $\{\sigma_t^2\}$.*

The condition $\max\{\alpha^{(1)}, \alpha^{(2)}\} + \max\{\beta^{(1)}, \beta^{(2)}\} < 1$ may rule out some commonly used cases, and we can relax it by revising the derivation for Claim (iii) in the proof of the theorem. However, the resulting condition have a complicated form since the distribution of ε_t is usually involved.

From Theorem 1, we can expect the geometric ergodicity of the buffered threshold GARCH process $\{y_t\}$. In the meanwhile, like all threshold models, strictly stationary solutions of the model defined in (2.1) and (2.2) have no closed form. We may want to extend the results of Theorem 1 to a more general buffered threshold model, but it seems impossible to define a suitable Markov chain related to the sequence of conditional variances $\{\sigma_t^2\}$ when $\max\{p, q, d\} > 1$.

Consider a general s-regime buffered threshold GARCH model with $s > 2$,

$$y_t = \varepsilon_t \sigma_t, \quad \sigma_t^2 = \omega^{(l)} + \sum_{i=1}^{q} \alpha_i^{(l)} y_{t-i}^2 + \sum_{j=1}^{p} \beta_j^{(l)} \sigma_{t-j}^2 \quad \text{if } R_t = l,$$

where $l = 1, \ldots, s$. There is more than one buffer zone, and this makes the definition of the regime-switching mechanism not unique. For example, there is more than one way to determine the regime of y_t when the threshold variable y_{t-d} jumps from one buffer zone into another. To make regime switching unambiguous, we proceed as follows. Let $-\infty = r_0 < r_1^L \leq r_1^U < r_2^L \leq r_2^U < \cdots < r_{s-1}^L \leq r_{s-1}^U < r_s = \infty$, so there are $s-1$ buffer zones, $(r_i^L, r_i^U]$ with $i = 1, \ldots, s-1$. The regime indicator R_t is equal to j if $r_{j-1}^U < y_{t-d} \leq r_j^L$, and the stochastic process stays in the same regime at time $t+1$ if the threshold variable y_{t-d+1} increases to the buffer zone $(r_j^L, r_j^U]$ or decreases to $(r_{j-1}^L, r_{j-1}^U]$. When the threshold variable y_{t-d+1} falls into the buffer zone $(r_i^L, r_i^U]$ with $i > j$ or $i < j-1$, the regime indicator R_{t+1} is set to i or $i+1$, respectively. By a method similar to Theorem 1, we can discuss the geometric ergodicity of its conditional variances for the case with $p = q = d = 1$.

3. Quasi-maximum Likelihood Estimation

This section considers the Gaussian quasi-maximum likelihood estimation (QMLE) for the two-regime buffered threshold GARCH model defined at (2.1) and (2.2).

Denote by $\boldsymbol{\lambda} = (\boldsymbol{\theta}', r_L, r_U, d)'$ the parameter vector of models (2.1) and (2.2), where $\boldsymbol{\theta}^{(1)} = (\omega^{(1)}, \alpha_1^{(1)}, \ldots, \alpha_q^{(1)}, \beta_1^{(1)}, \ldots, \beta_p^{(1)})'$, $\boldsymbol{\theta}^{(2)} = (\omega^{(2)}, \alpha_1^{(2)}, \ldots, \alpha_q^{(2)}, \beta_1^{(2)}, \ldots, \beta_p^{(2)})'$, and $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(1)\prime}, \boldsymbol{\theta}^{(2)\prime})'$. Let $\boldsymbol{\Theta}$ be a compact subset of $\mathbb{R}^{2p+2q+2}$, $[a, b]$ be a predetermined interval, and $d_{\max}$ be a predetermined positive integer. It is assumed that $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $a \leq r_L \leq r_U \leq b$, and $d \in D = \{1, \ldots, d_{\max}\}$. The true parameter vector is denoted by $\boldsymbol{\lambda}_0 = (\boldsymbol{\theta}_0', r_{0L}, r_{0U}, d_0)'$, where $\boldsymbol{\theta}_0^{(1)} = (\omega_0^{(1)}, \alpha_{01}^{(1)}, \ldots, \alpha_{0q}^{(1)}, \beta_{01}^{(1)}, \ldots, \beta_{0p}^{(1)})'$, $\boldsymbol{\theta}_0^{(2)} = (\omega_0^{(2)}, \alpha_{01}^{(2)}, \ldots, \alpha_{0q}^{(2)}, \beta_{01}^{(2)}, \ldots, \beta_{0p}^{(2)})'$, and $\boldsymbol{\theta}_0 = (\boldsymbol{\theta}_0^{(1)\prime}, \boldsymbol{\theta}_0^{(2)\prime})'$.

Define the regime indicator function as

$$R_t(r_L, r_U, d) = I(y_{t-d} \leq r_L) + I(r_L < y_{t-d} \leq r_U) R_{t-1}(r_L, r_U, d),$$

and the conditional variance function as

$$\sigma_t^2(\boldsymbol{\lambda}) = \Big[\omega^{(1)} + \sum_{i=1}^{q}\alpha_i^{(1)}y_{t-i}^2 + \sum_{j=1}^{p}\beta_j^{(1)}\sigma_{t-j}^2(\boldsymbol{\lambda})\Big]R_t(r_L, r_U, d)$$
$$+\Big[\omega^{(2)} + \sum_{i=1}^{q}\alpha_i^{(2)}y_{t-i}^2 + \sum_{j=1}^{p}\beta_j^{(2)}\sigma_{t-j}^2(\boldsymbol{\lambda})\Big][1 - R_t(r_L, r_U, d)]. \quad (3.1)$$

Then $R_t = R_t(r_{0L}, r_{0U}, d_0)$ and $\sigma_t^2 = \sigma_t^2(\boldsymbol{\lambda}_0)$. By temporarily assuming that ε_t is standard normal, we have the conditional log likelihood function of models (2.1) and (2.2), $-0.5L_n(\boldsymbol{\lambda}) - \log\sqrt{2\pi}$, where

$$L_n(\boldsymbol{\lambda}) = \sum_{t=1}^{n} l_t(\boldsymbol{\lambda}) \quad \text{and} \quad l_t(\boldsymbol{\lambda}) = \frac{y_t^2}{\sigma_t^2(\boldsymbol{\lambda})} + \log[\sigma_t^2(\boldsymbol{\lambda})].$$

Let $n_0 = \max\{q, d_{\max}\}$. For the observed time series $\{y_t, -n_0 + 1 \le t \le n\}$ generated by (2.1) and (2.2), the likelihood functions in the above depend on past observations infinitely far away, and hence initial values are needed.

For fixed r_L, r_U, and d, the first few observations of the threshold variable y_{t-d}, say $1 \le t \le t_0$, may fall into the buffer zone $(r_L, r_U]$ such that we fail to identify their regimes. These t_0 observations belong to the same regime since the threshold variable keeps staying at the buffer zone. We can simply assign them to regime one, and denote the resulting regime indicator function by $\widetilde{R}_t(r_L, r_U, d)$. We know the exact value of $R_{t_0+1}(r_L, r_U, d)$ since y_{t_0+1-d} is outside the buffer zone, and then it can be verified that $\widetilde{R}_t(r_L, r_U, d) = R_t(r_L, r_U, d)$ as $t_0 < t \le n$.

To evaluate the function $\sigma_t^2(\boldsymbol{\lambda})$, in addition to $\widetilde{R}_t(r_L, r_U, d)$ with $1 \le t \le t_0$, we also need initial values for $\sigma_{1-p}^2(\boldsymbol{\lambda}), \ldots, \sigma_0^2(\boldsymbol{\lambda})$; they can be set to $n^{-1}\sum_{t=1}^{n} y_t^2$ for simplicity. Accordingly, we denote by $\widetilde{\sigma}_t^2(\boldsymbol{\lambda})$, $\widetilde{l}_t(\boldsymbol{\lambda})$, and $\widetilde{L}_n(\boldsymbol{\lambda})$ the corresponding functions with these initial values, respectively. As a result, the Gaussian QMLE can be defined as

$$\widetilde{\boldsymbol{\lambda}}_n = (\widetilde{\boldsymbol{\theta}}_n', \widetilde{r}_L, \widetilde{r}_U, \widetilde{d})' = \operatorname*{argmin}_{\boldsymbol{\theta}\in\boldsymbol{\Theta}, d\in D, a\le r_L\le r_U\le b} \widetilde{L}_n(\boldsymbol{\theta}, r_L, r_U, d),$$

where $\boldsymbol{\lambda} = (\boldsymbol{\theta}', r_L, r_U, d)'$.

Assumption 1. *For $l = 1$ and 2, $\omega^{(l)} > 0$, $\alpha_i^{(l)} \ge 0$ with $1 \le i \le q$ and $\sum_{i=1}^{q}\alpha_i^{(l)} > 0$, $\beta_j^{(l)} \ge 0$ with $1 \le j \le p$ and $\sum_{j=1}^{p}\beta_j^{(l)} < 1$, and polynomials $1 - \sum_{j=1}^{p}\beta_j^{(l)}x^j$ and $\sum_{i=1}^{q}\alpha_i^{(l)}x^i$ have no common root.*

Assumption 2. *$\boldsymbol{\theta}_0^{(1)} \ne \boldsymbol{\theta}_0^{(2)}$, $P(y_t < a)\cdot P(y_t > b) > 0$, and ε_t has a bounded, continuous and positive density on $\mathbb{R}$.*

Theorem 2. *Suppose that the strictly stationary and ergodic time series* $\{y_t\}$ *is generated by the model at* (2.1) *and* (2.2), *with* $E|y_t|^{4+\delta} < \infty$ *for a small* $\delta > 0$. *If Assumptions* 1 *and* 2 *hold, then* $\widetilde{\boldsymbol{\lambda}}_n \to \boldsymbol{\lambda}_0$ *with probability one as* $n \to \infty$.

The moment condition $E|y_t|^{4+\delta} < \infty$ is mainly involved in deriving Claim (iii) in the proof of the theorem, and can be reduced to $E|y_t|^{2+\delta} < \infty$ if we further assume that the $\alpha_i^{(l)}$s and $\beta_j^{(l)}$ are bounded away from zero.

From Theorem 2, when the sample size n is large enough, the estimated delay parameter $\widetilde{d}$ is the true value d_0 since it only takes on integer values. Moreover, the estimated threshold parameters are usually super-consistent for the traditional threshold models as well as the buffered threshold models (Tong (1990); Li et al. (2015)), and it is then expected to hold for the buffered threshold GARCH model defined as in (2.1) and (2.2), $n(\widetilde{r}_L - r_{0L}) = O_p(1)$ and $n(\widetilde{r}_U - r_{0U}) = O_p(1)$. We leave the proofs of the super-consistency and the asymptotic distribution of the estimated threshold parameters $\widetilde{r}_L$ and $\widetilde{r}_U$ for future research. Without loss of generality, we assume that the values of (r_{0L}, r_{0U}, d_0) are known in deriving the asymptotic distribution of the QMLE, and then the parameter vector is $\boldsymbol{\theta}$.

Theorem 3. *Suppose the conditions in Theorem* 2 *hold. If* $E(\varepsilon_t^4) < \infty$ *and the true parameter vector* $\boldsymbol{\theta}_0$ *is an interior point of* $\boldsymbol{\Theta}$, *then* $\sqrt{n}(\widetilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \to_d N\{0, [E(\varepsilon_t^4) - 1]\Omega^{-1}\}$, *where*

$$\Omega = E\left(\frac{1}{\sigma_t^4(\boldsymbol{\theta}_0)}\frac{\partial \sigma_t^2(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}}\frac{\partial \sigma_t^2(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right).$$

This result can be proved by following the standard arguments for the asymptotic normality, and its proof is hence omitted. In practice, we can estimate the quantities $E(\varepsilon_t^4)$ and Ω respectively by

$$\frac{1}{n}\sum_{t=1}^n \frac{y_t^4}{\widetilde{\sigma}_t^4(\widetilde{\boldsymbol{\lambda}}_n)} \quad \text{and} \quad \widehat{\Omega}_n = \frac{1}{n}\sum_{t=1}^n \frac{1}{\widetilde{\sigma}_t^4(\widetilde{\boldsymbol{\lambda}}_n)}\frac{\partial \widetilde{\sigma}_t^2(\widetilde{\boldsymbol{\lambda}}_n)}{\partial \boldsymbol{\theta}}\frac{\partial \widetilde{\sigma}_t^2(\widetilde{\boldsymbol{\lambda}}_n)}{\partial \boldsymbol{\theta}'},$$

and hence the asymptotic variance in Theorem 3. It can be verified that they are all consistent. The initial values of the indicator function $\widetilde{R}_t(r_L, r_U, d)$ are equal to one. We can alternatively consider using zero for these initial values, and denote the corresponding functions by $\widetilde{R}_t^*(r_L, r_U, d)$ and $\widetilde{L}_n^*(\boldsymbol{\lambda})$, respectively. From proofs of Theorems 2 and 3, the effect of these initial values can be ignored asymptotically. As a result, we can define the QMLE $\widetilde{\boldsymbol{\lambda}}_n$ as $\text{argmin}_{\boldsymbol{\lambda}} \widetilde{L}_n^*(\boldsymbol{\lambda})$ or, even more precisely, $\text{argmin}_{\boldsymbol{\lambda}}\{\widetilde{L}_n(\boldsymbol{\lambda}), \widetilde{L}_n^*(\boldsymbol{\lambda})\}$.

When searching for the QMLE $\widetilde{\boldsymbol{\lambda}}_n$, we can first maximize $\widetilde{L}_n(\boldsymbol{\lambda})$ for each fixed (r_L, r_U, d),

$$\widehat{\boldsymbol{\theta}}_n(r_L, r_U, d) = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\text{argmin}}\ \widetilde{L}_n(\boldsymbol{\theta}, r_L, r_U, d),$$

and some commonly used computing algorithms such as Newton-Raphson can be employed to do the optimization. As in traditional threshold models (Li and Li (2011)), for an observed time series, the function $\widetilde{L}_n(\widehat{\boldsymbol{\theta}}_n(r_L, r_U, d), r_L, r_U, d)$ is piecewise constant with respect to r_L, r_U, and d, and has possible jumps at $\{y_{1-d}, \ldots, y_{n-d}\}$. Then we can further search for the estimators of (r_L, r_U, d) as,

$$(\widetilde{r}_L, \widetilde{r}_U, \widetilde{d}) = \underset{d\in D, a\le r_L\le r_U\le b}{\operatorname{argmin}} \widetilde{L}_n(\widehat{\boldsymbol{\theta}}_n(r_L, r_U, d), r_L, r_U, d),$$

where r_L and r_U take values on $\{y_{1-d}, \ldots, y_{n-d}\} \cap [a, b]$. In practice, we can choose the values of a and b to be some empirical percentiles of $\{y_t\}$; see Ling and Tong (2005), Li and Li (2008), and Zhu, Yu, and Li (2014). It can be verified that $\widetilde{\boldsymbol{\theta}}_n = \widehat{\boldsymbol{\theta}}_n(\widetilde{r}_L, \widetilde{r}_U, \widetilde{d})$. When the sample size n is large, searching for $(\widetilde{r}_L, \widetilde{r}_U, \widetilde{d})$ is time-consuming, and we can alternatively consider a grid searching algorithm for r_L and r_U.

4. Simulation Studies

This section reports on two simulation experiments to study the finite sample performance of the Gaussian QMLE in the previous section. For each generated sequence, the range of boundary parameters r_L and r_U was set to from the 25th to the 75th empirical percentiles, and the maximum of the delay parameter d was six.

The data generating process in the first simulation experiment was a buffered threshold ARCH process,

$$y_t = \varepsilon_t \sigma_t, \qquad \sigma_t^2 = \begin{cases} \omega^{(1)} + \sum_{i=1}^{7} \alpha_i^{(1)} y_{t-i}^2, & \text{if } R_t = 1, \\ \omega^{(2)} + \sum_{i=1}^{7} \alpha_i^{(2)} y_{t-i}^2, & \text{if } R_t = 0, \end{cases}$$

with the regime indicator

$$R_t = \begin{cases} 1, & \text{if } y_{t-4} \le -0.05, \\ R_{t-1}, & \text{if } -0.05 < y_{t-4} \le 0.08, \\ 0, & \text{if } y_{t-4} > 0.08, \end{cases}$$

where $\{\varepsilon_t\}$ were independent and normally distributed with mean zero and variance one. The sample size was set to n =4,000, and the grid searching algorithm was employed for r_L and r_U to save computation time. There were 350 replications calculated. The QMLE $\widetilde{\boldsymbol{\lambda}}_n$ is summarized in Table 1, and the delay parameter d is correctly identified for 340 out of 350 replications. We tried larger sample sizes, and could observe decrease of the bias and the empirical standard deviations. Figure 1 gives the histograms of $\sqrt{n}(\widetilde{\omega}_n^{(l)} - \omega_0^{(l)})$ with $l = 1$ and 2, and the large sample normality result is then confirmed. A similar finding is observed from histograms of the other parameters in $\boldsymbol{\theta}$, and hence they are omitted to save

Table 1. Simulation results for buffered threshold ARCH models.

	Lower regime			Upper regime				Other parameters		
	$\boldsymbol{\lambda}_0$	Bias	ESD	$\boldsymbol{\lambda}_0$	Bias	ESD		$\boldsymbol{\lambda}_0$	Bias	ESD
ω	0.02	0.0003	0.0022	0.01	-0.0004	0.0018	r_L	-0.05	-0.0004	0.0286
α_1	0.20	0.0009	0.0387	0.25	0.0031	0.0461	r_U	0.08	0.0058	0.0328
α_2	0.30	0.0010	0.0446	0.20	-0.0001	0.0430	d	4	-0.0225	0.3181
α_3	0.04	0.0036	0.0290	0.10	-0.0016	0.0295				
α_4	0.08	-0.0036	0.0300	0.08	0.0038	0.0302				
α_5	0.05	-0.0029	0.0284	0.09	-0.0002	0.0314				
α_6	0.07	-0.0025	0.0311	0.10	-0.0022	0.0291				
α_7	0.03	0.0005	0.0224	0.01	0.0055	0.0170				

$\boldsymbol{\lambda}_0$: true parameters; ESD: empirical standard deviation.

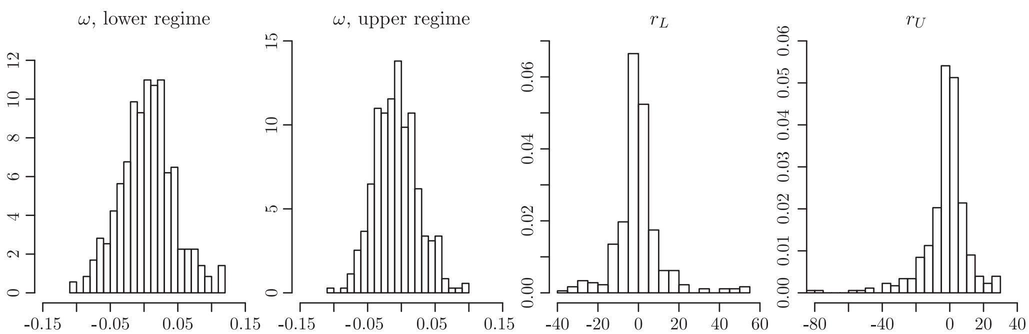

Figure 1. Histograms of $\sqrt{n}(\widetilde{\omega}_n^{(1)} - \omega_0^{(1)})$, $\sqrt{n}(\widetilde{\omega}_n^{(2)} - \omega_0^{(2)})$, $n(\widetilde{r}_L - r_{0L})$ and $n(\widetilde{r}_U - r_{0U})$ for buffered threshold ARCH models (from left to right).

space. Histograms of $n(\widetilde{r}_L - r_{0L})$ and $n(\widetilde{r}_U - r_{0U})$ are also presented in Figure 1. It can be seen that their shapes are much more peaked than that of the normal distribution, and resemble those of the two-sided compound poisson processes in Li, Ling, and Li (2013).

In the second experiment, we generated time series by the buffered threshold GARCH model,

$$y_t = \varepsilon_t \sigma_t, \qquad \sigma_t^2 = \begin{cases} \omega^{(1)} + \alpha_1^{(1)} y_{t-1}^2 + \beta_1^{(1)} \sigma_{t-1}^2, & \text{if } R_t = 1, \\ \omega^{(2)} + \alpha_1^{(2)} y_{t-1}^2 + \beta_1^{(2)} \sigma_{t-1}^2, & \text{if } R_t = 0, \end{cases}$$

where $\{\varepsilon_t\}$ and regime indicators $\{R_t\}$ were the same as that in the first experiment except that $d_0 = 3$. We set the sample size to n =3,600. The grid searching algorithm was employed again for r_L and r_U, and there were 200 replications. The QMLE $\widetilde{\boldsymbol{\lambda}}_n$ is summarized in Table 2, and the delay parameter d is correctly identified for 124 out of 200 replications. Figure 2 gives the histograms of $\sqrt{n}(\widetilde{\alpha}_{1n}^{(l)} - \alpha_{01}^{(l)})$ with $l = 1$ and 2, and normality can be expected again. Histograms of the other parameters in $\boldsymbol{\theta}$ have similar patterns, and hence

Table 2. Simulation results for buffered threshold GARCH models.

	Lower regime			Upper regime				Other parameters		
	$\boldsymbol{\lambda}_0$	Bias	ESD	$\boldsymbol{\lambda}_0$	Bias	ESD		$\boldsymbol{\lambda}_0$	Bias	ESD
ω	0.02	0.0028	0.0224	0.05	0.0017	0.0256	r_L	-0.05	-0.0195	0.1177
α_1	0.06	0.0038	0.0298	0.10	0.0080	0.0333	r_U	0.08	0.0063	0.1111
β_1	0.80	0.0227	0.1170	0.70	0.0040	0.1308	d	3	0.0493	1.0889

$\boldsymbol{\lambda}_0$: true parameters; ESD: empirical standard deviation.

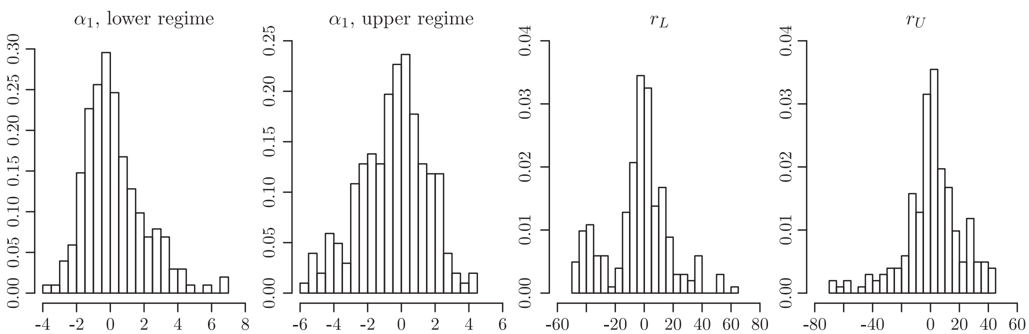

Figure 2. Histograms of $\sqrt{n}(\widetilde{\alpha}_{1n}^{(1)} - \alpha_{01}^{(1)})$, $\sqrt{n}(\widetilde{\alpha}_{1n}^{(2)} - \alpha_{01}^{(2)})$, $n(\widetilde{r}_L - r_{0L})$ and $n(\widetilde{r}_U - r_{0U})$ for buffered threshold GARCH models (from left to right).

are omitted here. Histograms of $n(\widetilde{r}_L - r_{0L})$ and $n(\widetilde{r}_U - r_{0U})$ are presented in Figure 2. These are similar to those of the buffered ARCH models in Figure 1. As in the classical case, it appears that the buffered GARCH model is more difficult to fit than the pure buffered ARCH models.

5. An Empirical Example

This section considers the daily closing prices, adjusted for dividends and splits, of Honeywell International Inc (HON), one of the components of the Dow Jones Industrial Average index. We focus on the sequence of log returns in percentage from June 11, 1990 to September 12, 2006, and there are 4,099 observations in total. This time series has been studied by Caiado and Crato (2010), and was shown by the Ljung-Box test to have no significant serial correlation effect. Figure 3 gives the sample autocorrelation functions (ACFs) of log returns and squared log returns, and a pure volatility model is then suggested.

We first applied the buffered GARCH model to the sequence, and the estimating procedure in Section 3 was employed to search for the estimates with orders p and q being fixed at one. The range of boundary parameters r_L and r_U was from the 15th to the 85th empirical percentiles of observations, and the

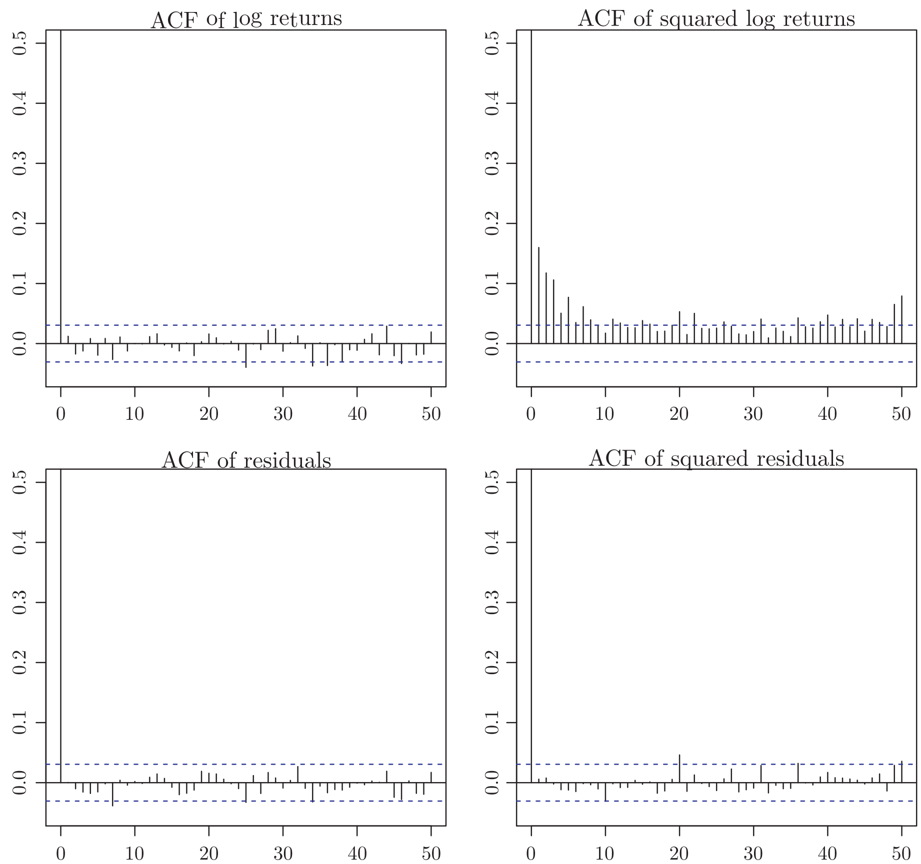

Figure 3. Sample ACFs of log returns and squared log returns for daily closing prices of HON (upper panel), and sample ACFs of residuals and squared residuals from the fitted buffered threshold GARCH model (lower panel).

delay parameter was from one to six. The fitted model had the form

$$y_t = \varepsilon_t \sigma_t, \quad \sigma_t^2 = \begin{cases} 0.2840_{0.1294} + 0.1647_{0.0374} y_{t-1}^2 + 0.8353_{0.0504} \sigma_{t-1}^2, & \text{if } R_t = 1, \\ 0.1384_{0.0735} + 0.0772_{0.0242} y_{t-1}^2 + 0.8366_{0.0377} \sigma_{t-1}^2, & \text{if } R_t = 0, \end{cases}$$

with the regime indicator

$$R_t = \begin{cases} 1, & \text{if } y_{t-1} \le -1.3209, \\ R_{t-1}, & \text{if } -1.3209 < y_{t-1} \le 1.3212, \\ 0, & \text{if } y_{t-1} > 1.3212, \end{cases}$$

where the subscripts of parameter estimates are their associated standard errors, with the Bayesian information criterion (BIC) of 16,816. The ACFs of residuals and squared residuals are presented in Figure 3, and they slightly stand out from the 95% confidence limits at only a few lags. As a result, we conclude the adequacy of the fitted model. To further evaluate adequacy, the ad hoc McLeod-Li test statistic (McLeod and Li (1983)) was calculated for the squared standardized residuals, with $Q(20) = 18.53$, $Q(50) = 47.99$ and $Q(70) = 84.61$, and corresponding p-values all greater than 0.10.

The three-regime threshold GARCH model also has two threshold parameters, and the buffered threshold model includes the two-regime threshold model as a special case. As a comparison, it is of interest to further fit the three-regime threshold GARCH model to the data. The settings were the same as those for the aforementioned buffered threshold model, and the fitted model was

$$y_t = \varepsilon_t \sigma_t,$$

$$\sigma_t^2 = \begin{cases} 0.3796_{0.2931} + 0.1838_{0.0455} y_{t-1}^2 + 0.8162_{0.0939} \sigma_{t-1}^2, & \text{if } y_{t-1} \leq -1.3209, \\ 0.0461_{0.0867} + 0.0925_{0.1228} y_{t-1}^2 + 0.9075_{0.0363} \sigma_{t-1}^2, & \text{if } -1.3209 < y_{t-1} \leq 1.3812, \\ 0.1346_{0.2722} + 0.0905_{0.0279} y_{t-1}^2 + 0.7291_{0.0803} \sigma_{t-1}^2, & \text{if } y_{t-1} > 1.3812, \end{cases}$$

with a BIC of 16,835, somewhat larger than that of the fitted buffered threshold model. The McLeod-Li test statistic had the value of $Q(20) = 21.58$, $Q(50) = 54.92$ and $Q(70) = 89.79$, are all greater than the corresponding values for the fitted buffered threshold model. As the McLeod-Li test is based on the autocorrelations of squared residuals, we conclude that the buffered threshold model has a better performance in interpreting the squared log returns, or simply its volatility. Especially, the p-value of $Q(70)$ is even smaller than 0.10, and the fitted three-regime threshold model may not be adequate.

The fitted buffer zone is almost the same as the middle regime of the fitted three-regime threshold model, and the zone includes roughly 60% of the observations. We can argue that these observations have the same structure as in the lower or upper regime rather than following a separate one. Moreover, in the fitted buffered threshold model, the time series exhibits a stronger persistence at the lower regime, and it can be interpreted as that bad news may have longer influence to the volatility.

6. Conclusions

This paper proposes a new type of threshold conditional heteroscedastic model with the regime-switching mechanism possessing a buffered region so that the regime-switching is delayed. This is illustrated with a US stock example. This nonlinear model can provide some new insight into the asymmetric behavior of volatilities of financial time series.

This paper also derives some basic results of Gaussian maximum likelihood estimation, including strong consistency and asymptotic normality. However, like the beginning of many statistical models, there remain many open problems, such as how to construct tests to check whether the buffered GARCH model can provide a better fit compared with the common GARCH model or even with the traditional threshold GARCH model. We look to these technical problems in future research.

Acknowledgements

The authors would like to thank the Editor, an associate editor and two anonymous referees for valuable comments that led to the substantial improvement of this paper, and Hong Kong RGC grant HKU 703711P for partial support. We would like to thank Wenxin Cai and Renjie Lu for their help in data analysis. W. K. Li greatly appreciated the invitation to present the first version of this paper in the conference, Frontiers of Statistics and Forecasting in celebrating the 80th birthday of Professor George C. Tiao, December 16-18, 2013, Academia Sinica, Taipei.

Supplementary Materials

The online supplementary material includes the proofs of Theorems 1 and 2.

References

Bekaert, G. and Wu, G. (2000). Asymmetric volatility and risk in equity markets. *Rev. Finan. Studies* **13**, 1-42.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**, 307-327.

Brooks, C. (2001). A double-threshold GARCH model for the French Franc/Deutschmark exchange rate. *J. Forecasting* **20**, 135-143.

Caiado, J. and Crato, N. (2010). Identifying common dynamic features in stock returns. *Quantitative Finance* **10**, 797-807.

Chan, K. S. (1993). Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Ann. Statist.* **21**, 520-533.

Chan, K. S. and Tong, H. (1986). On estimating thresholds in autoregressive models. *J. Time Series Anal.* **7**, 178-190.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987-1007.

Francq, C. and Zakoïan, J. (2006). Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econom. Theory* **22**, 815-834.

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

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica* **68**, 575-603.

Li, D., Ling, S. and Li, W. K. (2013). Asymptotic theory on the least squares estimation of threshold moving-average models. *Econom. Theory* **29**, 482-516.

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. *Statist. 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, W. K. and Li, C. W. (1996). On a double-threshold autoregressive heteroscedastic time series model. *J. Appl. Econom.* **11**, 253-274.

Ling, S. and Tong, H. (2005). Testing for a linear MA model against threshold MA models. *Ann. Statist.* **33**, 2529-2552.

Liu, J., Li, W. K. and Li, C. W. (1997). On a threshold autoregression with conditional heteroscedastic variances. *J. Statist. Plann. Inference* **62**, 279-300.

McCulloch, R. and Tsay, R. S. (1994). Statistical analysis of economic time series via markov switching models. *J. Time Series Anal.* **15**, 235-250.

McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared residual autocorrelations. *J. Time Series Anal.* **4**, 269-273.

Meitz, M. and Saikkonen, P. (2008). Ergodicity, mixing, and existence of moments of a class of markov models with applications to GARCH and ACD models. *Econom. Theory* **24**, 1291-1320.

Tong, H. (1978). On a threshold model. In *Pattern Recognition and Signal Processing* (Edited by C. H. Chen), 575-586, Sijthoff and Noordhoff, Amsterdam.

Tong, H. (1990). *Nonlinear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford.

Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *J. Roy. Statist. Soc. Ser. B* **42**, 245-292.

Tsay, R. S. (1998). Testing and modeling multivariate threshold models. *J. Amer. Statist. Assoc.* **93**, 1188-1202.

van Dijk, D., Terasvirta, T. and Franses, P. H. (2002). Smooth transition autoregressive models - a survey of recent developments. *Econom. Rev.* **21**, 1-47.

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

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

Department of Statistics and Actuarial Science, University of Hong Kong, Hong Kong.
E-mail: hanglph@gmail.com

Department of Statistics and Actuarial Science, University of Hong Kong, Hong Kong.
E-mail: hrntlwk@hku.hk

Department of Statistics and Actuarial Science, University of Hong Kong, Hong Kong.
E-mail: plhyu@hku.hk

Department of Statistics and Actuarial Science, University of Hong Kong, Hong Kong.
E-mail: gdli@hku.hk

(Received March 2014; accepted October 2015)