JOURNAL OF TIME SERIES ANALYSIS
J. Time. Ser. Anal. **36:** 26–38 (2015)
Published online 7 October 2014 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1111/jtsa.12089

ORIGINAL ARTICLE

HYPOTHESIS TESTING FOR ARCH MODELS: A MULTIPLE QUANTILE REGRESSIONS APPROACH

SEONJIN KIM*
Department of Statistics, Miami University, Oxford, OH 45056, USA

We propose a quantile regression-based test to detect the presence of autoregressive conditional heteroscedasticity by combining distributional information across multiple quantiles. A chi-square-type test statistic based on the weighted average of distinct regression quantile estimators is formed. Unlike the widely used likelihood-based tests, the proposed test does not make any distributional assumptions on the underlying errors. Monte Carlo simulation studies show that the proposed test outperforms the likelihood-based tests in several aspects.

Received 24 February 2014; Revised 8 May 2014; Accepted 8 July 2014

Keywords: ARCH model; conditional heteroscedasticity; hypothesis testing; quantile regression.

1. INTRODUCTION

The autoregressive conditional heteroscedasticity (ARCH) model has been widely used to model the volatility of economic and financial time series data since its introduction by Engle (1982). The ARCH model and its generalizations, especially the generalized ARCH (GARCH) model (Bollerslev, 1986), provide an appealing structure for the theory that the current volatility is determined by past observations. Furthermore, it is well known that neglecting ARCH effects lead to some disadvantages including loss in asymptotic efficiency of parameter estimation (Engle, 1982) and overparameterization of an autoregressive moving average model (Weiss, 1984). Thus, so as to take advantage of the interpretability of the ARCH model and avoid the aforementioned disadvantages of neglecting ARCH effects, it is necessary to test for the existence of conditional heteroscedasticity in time series modelling.

Much effort has been devoted to developing tests to verify ARCH effects: the Lagrange multiplier (LM) test (Engle, 1982), the locally most mean powerful based score (LBS) test exploiting the one-sided nature of the null hypothesis (Lee and King, 1993), a test for ARCH effects in the frequency domain (Hong and Shehadeh, 1999) and Monte Carlo simulation-based finite-sample tests (Dufour *et al.*, 2004). Many of the existing tests of ARCH effects are based on the likelihood function of the errors, so they require distributional assumptions. For example, the LM test statistic (Engle, 1982) has a chi-square asymptotic distribution with p degrees of freedom when the errors are normally distributed. However, it should be emphasized that the normality assumption is naturally incorrect in financial return data that the ARCH process is originally designed to model. As is well known, these data typically follow an asymmetric and heavy-tailed distribution reflecting traders' behaviours. Traders usually react more strongly to negative news than positive news and often react extremely to an event. Nevertheless, the LM test has been widely used for non-normal cases, but the asymptotic properties of the statistic are not clear when the errors are not normally distributed.

* Correspondence to: Seonjin Kim, Department of Statistics, Miami University, 334D Upham Hall, Oxford, OH 45056, USA. E-mail: kims20@miamioh.edu

To alleviate the aforementioned issue, a quantile regression (QR)-based method is proposed. QR-based approaches explore the distribution of the errors, while likelihood-based methods assume a distribution. Despite this attractive feature, little has been done to develop tests for ARCH effects using QR. Koenker and Zhao (1996) mentioned inference based on stacking several distinct regression quantile (RQ) estimators but did not construct an ARCH effects test. Furno (2004) proposed the RQ test to verify the presence of heteroscedasticity when comparing the slope parameters of the regressions computed at different quantiles.

Recently, there has been a growing interest in studying QR-based estimation combining information across multiple quantiles under various frameworks. Some of these works include simple linear models (Koenker, 1984; Portnoy and Koenker, 1989; Zou and Yuan, 2008), the GARCH model (Xiao and Koenker, 2009), non-parametric regression models (Kai *et al.*, 2010), a broad range of regression models (Zhao and Xiao, 2014) and time-varying coefficient longitudinal models (Kim *et al.*, 2014) among others. These references reveal that the estimates constructed by properly aggregating information across multiple quantiles are more efficient than the least squares estimate when the errors are not normally distributed and they convey almost equivalent performance when the errors are normal.

Motivated by the fact that a combination of multiple quantiles can afford a more complete knowledge of the unknown distribution, a powerful and distribution-free QR-based test is established. First, we address the asymptotic normality of the RQ estimator at a single quantile and form a chi-square-type ARCH effects test at any quantile under the null hypothesis of no ARCH effects. Then, we propose an idea to combine information across multiple quantiles. The proposed test called the weighted average quantile (WAQ) test employs a weighted average of the RQ estimators at multiple quantiles to derive a test statistic. For comparison, another QR-based test motivated by the idea of Koenker and Zhao (1996), which stacks several RQ estimators, is also constructed. Unlike likelihood-based tests, our proposed WAQ test does not impose any distributional assumptions, and moreover, it automatically makes use of the unknown distributional information. Simulation studies indicate that the proposed WAQ test outperforms the widely used LM test as well as the LBS test.

The remainder of this article is organized as follows. Section 2 introduces the QR-based test using a single quantile. In Section 3, we introduce two ARCH detection tests combining information across multiple quantiles. Section 4 presents simulation results. Proofs are provided in Section 5.

2. ARCH EFFECTS TEST BASED ON QUANTILE REGRESSION

In this article, we consider an ARCH model of order p given by

$$X_i = \sigma_i \varepsilon_i, \quad i = 1, \dots, n, \quad \text{where} \quad \sigma_i = \beta_0 + \beta_1 |X_{i-1}| + \cdots + \beta_p |X_{i-p}|, \tag{1}$$

where ε_i are i.i.d. random variables with $\mathbb{E}(\varepsilon_i) = 0$ and $\text{var}(\varepsilon_i) = 1$ and are independent of past observations $X_{i-1}, X_{i-2}, \dots$. The coefficients $\beta_0 > 0, 0 \leq \beta_j < 1$ for $j = 1, \dots, p$ satisfy $\sum_{j=1}^{p} \beta_j < 1$ to ensure a stationary solution. The structure of the conditional variance in (1) is slightly different from the original quadratic form of the ARCH model proposed by Engle (1982) given by

$$X_i = \sigma_i \varepsilon_i, \quad i = 1, \dots, n, \quad \text{where} \quad \sigma_i^2 = \beta_0 + \beta_1 X_{i-1}^2 + \cdots + \beta_p X_{i-p}^2. \tag{2}$$

As noted by Xiao and Koenker (2009), the linear ARCH model in (1) is less sensitive to extreme returns than the quadratic ARCH model in (2). Also, the linear structure is commonly employed in the literature on QR for the ARCH or GARCH model because of its suitable structure (Koenker and Zhao, 1996; Xiao and Koenker, 2009).

Our goal in this paper is to test the null hypothesis of no ARCH effects defined as

$$H_0 : \beta_1 = \cdots = \beta_p = 0, \quad \text{or equivalently} \quad \sigma_i = \beta_0 \text{ is constant.} \tag{3}$$

DOI: 10.1111/jtsa.12089

When the previous null hypothesis is tested, the choice of order p is not critical. The primary concern in testing ARCH effects is to detect the existence of conditional heteroscedasticity. Thus, once the null hypothesis in (3) is rejected, we can select an appropriate model using Akaike information criterion or Bayesian information criterion. For detailed discussions, see chapter 4.2.3 in the work of Fan and Yao (2005). In this sense, the proposed tests in this article are readily extended for testing GARCH effects because under a mild condition, a strictly stationary GARCH (p,q) is effectively an ARCH (∞).

So as to develop a hypothesis test of ARCH effects, we do not directly use the ARCH model in (1). Let $Y_i = |X_i|, e_i = |\varepsilon_i|, Z_i = (1, Y_{i-1}, \ldots, Y_{i-p})^T$ and $\beta = (\beta_0, \beta_1, \ldots, \beta_p)^T$. Taking the absolute value of both sides of (1), we have

$$Y_i = \sigma_i |\varepsilon_i| = (\beta_0 + \beta_1 Y_{i-1} + \cdots + \beta_p Y_{i-p})\, e_i = Z_i^T \beta e_i, \quad i = 1, \ldots, n.$$

Since all the coefficients and the components of Z_i are non-negative, σ_i is invariant under the transformation, and thus, the null hypothesis can be tested with the absolute value series $\{Y_1, Y_2, \ldots, Y_n\}$. Denote by $Q_e(\tau)$ the τ-quantile of e_i and by $Q_{Y_i|Z_i}(\tau)$ the conditional τ-quantile of Y_i given Z_i. Then, we can write

$$Q_{Y_i|Z_i}(\tau) = Z_i^T \beta Q_e(\tau) = Z_i^T \beta(\tau), \quad \text{where} \quad \beta(\tau) = Q_e(\tau)\beta.$$

Applying QR, we obtain the τth RQ estimator

$$\hat{\beta}(\tau) = \underset{b}{\operatorname{argmin}} \sum_i \rho_\tau \left(Y_i - Z_i^T b\right), \tag{4}$$

where $\rho_\tau(z) = z\{\tau - \mathbf{1}(z \leq 0)\}$ is the quantile loss function at a quantile $\tau \in (0,1)$ and $\mathbf{1}(\cdot)$ is an indicator function. In particular, $\tau = 0.5$ corresponds to the least absolute deviation estimator. It is worth noting that formulating the RQ estimator in terms of the absolute value $Y_i = |X_i|$ is not necessary to derive the asymptotic properties of $\hat{\beta}(\tau)$. Koenker and Zhao (1996) studied asymptotic properties of an RQ estimator $\tilde{\beta}(\tau) = \operatorname{argmin}_b \sum_i \rho_\tau \left(X_i - Z_i^T b\right)$ with the original response X_i. For $\tau \in (0,1), \hat{\beta}(\tau)$ and $\tilde{\beta}(\tau)$ are consistent estimators of $Q_e(\tau)\beta$ and $Q_\varepsilon(\tau)\beta$ respectively, where $Q_\varepsilon(\tau)$ is the τ-quantile of ε_i. In other words, they identify β up to a scaling factor but in fact estimate different quantities.

There are several nice features of $\hat{\beta}(\tau)$ over $\tilde{\beta}(\tau)$ for developing ARCH effects test. First, $\hat{\beta}(\tau)$ and $\tilde{\beta}(\tau)$ estimate $\mathbf{0}$ when either $\beta = \mathbf{0}$ or $Q_e(\tau) = 0$ and $Q_\varepsilon(\tau) = 0$ respectively, where $\mathbf{0}$ is a zero column vector. Without prior information about the underlying distribution, we do not know where $Q_\varepsilon(\tau_1) = 0$. Thus, extra work is necessary to find τ_1 to avoid power loss in hypothesis testing. In contrast, since $Q_e(0) = 0$ regardless of the type of distribution, any quantile $\tau \in (0,1)$ can be used for $\hat{\beta}(\tau)$ immediately. Second, when the density function of the errors ε_i is symmetric about 0, if $\tilde{\beta}(\tau)$ is employed to construct the proposed test statistic in Section 3.1, which aggregates distributional information across multiple quantiles, then the test statistic is a consistent estimator of 0 under both the null and alternative hypotheses. As a result, it leads to poor power in testing ARCH effects. See Remark 1, for more discussion. In this paper, we focus on testing ARCH effects rather than focusing on estimation of the parameters.

Throughout the rest of the article, for a random vector Z, we write $Z \in \mathcal{L}^q, q > 0$ if $\|Z\|_q := \{E(|X|^q)\}^{1/q} < \infty$. Consider the following regularity conditions.

Assumption 1.

(i) ε_i is independent of $\mathcal{F}_{i-1} = \sigma(\varepsilon_{i-1}, \varepsilon_{i-2}, \ldots)$ for all i.

(ii) Denote by F_e and f_e the distribution and density function of e_i respectively. f_e is positive, continuous and bounded on $\{u : 0 < F_e(u) < 1\}$.

DOI: 10.1111/jtsa.12089

(iii) $\{Y_i, Y_{i-1}, \ldots, Y_{i-p}\}$ is an α-mixing stationary process with $Y_i \in \mathcal{L}^\delta$ for some $\delta > 2$, and the mixing coefficient α_k satisfies $\sum_{k=1}^{\infty} \alpha_k^{1-2/\delta} < \infty$.
(iv) $\mathbb{E}\left(Z_1 Z_1^T\right)$ is positive definite.

The regularity conditions are mild and commonly imposed in time series analysis. Under these conditions, we can derive the asymptotic normality of $\hat{\beta}(\tau)$.

Theorem 1. Suppose that Assumption 1 holds. Define $V_k = \mathbb{E}\left(Z_1 Z_1^T / \sigma_1^k\right)$.

(i) Then, we have the asymptotic Bahadur representation

$$\hat{\beta}(\tau) - \beta(\tau) = \frac{V_1^{-1}}{n f_e(Q_e(\tau))} \sum_{i=1}^{n} Z_i \left\{\tau - \mathbf{1}\left(e_i < Q_e(\tau)\right)\right\} + o_p(1). \tag{5}$$

(ii) The following asymptotic normality holds

$$\sqrt{n}\left\{\hat{\beta}(\tau) - \beta(\tau)\right\} \Rightarrow N\left(\mathbf{0}, \frac{\tau(1-\tau)}{f_e^2(Q_e(\tau))} V_1^{-1} V_0 V_1^{-1}\right). \tag{6}$$

As discussed earlier, Theorem 1 demonstrates that for $\tau \in (0,1)$, $\hat{\beta}(\tau)$ is a consistent estimator of $\beta(\tau)$. Denote by $\beta^* = (\beta_1, \ldots, \beta_p)^T$ the vector of all elements of β, but the intercept β_0 and by $\hat{\beta}^*(\tau)$ the vector of its corresponding RQ estimator at a quantile τ. So as to test the null hypothesis of no ARCH effects, it is not necessary to estimate β directly because when $\beta^* = \mathbf{0}$ (i.e. when the null hypothesis is true), $\hat{\beta}^*(\tau)$ is a consistent estimator of $\mathbf{0}$ regardless of the choice of τ. In addition, $Q_e(\tau)$ is non-negative since $e_i = |\varepsilon_i|$. Thus, so as to test ARCH effects, it is sufficient to test the null hypothesis

$$H_0(\tau): \beta_1(\tau) = \cdots = \beta_p(\tau) = 0, \tag{7}$$

for any $\tau \in (0,1)$.

Denote by f_Y and $Q_Y(\tau)$ the density function of Y_i and the τ-quantile of Y_i respectively. Since $\sigma_i = \beta_0$ is constant under the null hypothesis, we have $f_Y(Q_Y(\tau)) = f_{Y/\sigma_i}(Q_{Y/\sigma_i}(\tau))/\sigma_i = f_e(Q_e(\tau))/\sigma_i$, and thus, we can simplify the limiting variance in (6).

Corollary 1. Suppose that Assumption 1 holds. Under the null hypothesis of no ARCH effects, we have

$$\sqrt{n}\left\{\hat{\beta}(\tau) - \left(\beta_0, \mathbf{0}^T\right)^T\right\} \Rightarrow N\left(0, \frac{\tau(1-\tau)V_0^{-1}}{f_Y^2(Q_Y(\tau))}\right). \tag{8}$$

The limiting variance of the null distribution in (8) is invariant under the value of the variance $\sigma_i = \beta_0$. Consequently, the proposed test statistic later is not affected by β_0. Intuitively, it is reasonable because the value of β_0 – that is, the variance of the errors under the null hypothesis – should not influence the result of an ARCH effects test. Based on the previous discussion, we can form a chi-square-type ARCH effects test at each quantile τ, as shown in Corollary 2.

DOI: 10.1111/jtsa.12089

Corollary 2. Suppose that Assumption 1 holds. Let W be the sub-element of V_0^{-1} corresponding to $\beta^*(\tau)$. Then, under the null hypothesis in (7), we have

$$T(\tau) := \frac{n f_Y^2(Q_Y(\tau))}{\tau(1-\tau)} \hat{\beta}^*(\tau)^T W^{-1} \hat{\beta}^*(\tau) \Rightarrow \chi_p^2, \tag{9}$$

where χ_p^2 is the chi-square distribution with p degrees of freedom.

A hypothesis test for ARCH effects can be performed with the test statistic $T(\tau)$ in (9). However, a problem with this QR-based test is that the value of the test statistic may change substantially with respect to the quantile τ and the distribution of Y_i. This instability may be caused by using partial distributional information, and thus, there is a need to develop a method integrating the partial information that each quantile contains.

3. ARCH EFFECTS TESTS ACROSS MULTIPLE QUANTILES

3.1. Weighted Average Quantile Test

As discussed in Section 2, each quantile involves partial information, and thus, it is desirable to aggregate distributional information over quantiles. To this end, we adopt the weighted average of multiple single RQ estimators at different τ's as an aggregate estimator. Without prior information, it is reasonable to choose uniformly spaced quantiles, but the derived results and proposed tests in this article can be employed with any sets of quantiles. Throughout the rest of this article, let $\tau_r = r/(t+1), r = 1, \ldots, t$, be t uniformly spaced quantiles and $w = (w_1, \ldots, w_t)^T$ be a vector of weights. Then, the WAQ estimator is defined as

$$\hat{\beta}_{\mathrm{QA}}(w) = \sum_{r=1}^{t} w_r \hat{\beta}(\tau_r), \quad \sum_{r=1}^{t} w_r = 1, \tag{10}$$

and the corresponding parameter $\beta_{\mathrm{QA}}(w) = \sum_{r=1}^{t} w_r \beta(\tau_r)$. Recall that β^* is the vector of all elements of β but the intercept β_0 and $\hat{\beta}^*(\tau)$ is its corresponding RQ estimate at a single quantile τ. Since each $\hat{\beta}(\tau_r)$ estimates $\beta(\tau_r)$ and not β, $\hat{\beta}_{\mathrm{QA}}(w)$ is not a consistent estimator of β. However, since all $\beta(\tau_r) = Q_e(\tau_r)\beta, r = 1, \ldots, t$ are proportional to β, all $\hat{\beta}^*(\tau_r)$ are estimating $\mathbf{0}$ when the null hypothesis of no ARCH effects is true. Motivated by this fact, we establish a new ARCH effects test combining information across multiple quantiles.

Theorem 2. Suppose that Assumption 1 holds. Define $\hat{\beta}_{\mathrm{QA}}^*(w) = \sum_{r=1}^{t} w_r \hat{\beta}^*(\tau_r)$.

(i) Then, the following asymptotic normality holds

$$\sqrt{n}\left\{\hat{\beta}_{\mathrm{QA}}(w) - \beta_{\mathrm{QA}}(w)\right\} \Rightarrow N\left(\mathbf{0}, s V_1^{-1} V_0 V_1^{-1}\right), \tag{11}$$

where $s = \sum_{r=1}^{t} \sum_{r'=1}^{t} \frac{\min(\tau_r, \tau_{r'}) - \tau_r \tau_{r'}}{f_e(Q_e(\tau_r)) f_e(Q_e(\tau_{r'}))} w_r w_{r'}$.

(ii) Under the null hypothesis in (3), we have

$$\sqrt{n}\left\{\hat{\beta}_{\mathrm{QA}}(w) - \left(\sum_{r=1}^{t} w_r \beta_0(\tau_r), \mathbf{0}^T\right)^T\right\} \Rightarrow N\left(0, s' V_0^{-1}\right), \tag{12}$$

DOI: 10.1111/jtsa.12089

where $s' = \sum_{r=1}^{t} \sum_{r'=1}^{t} \frac{\min(\tau_r, \tau_{r'}) - \tau_r \tau_{r'}}{f_Y(Q_Y(\tau_r)) f_Y(Q_Y(\tau_{r'}))} w_r w_{r'}$.

(iii) Recall W in Corollary 2. Under the null hypothesis in (3), we have

$$T_{QA}(w) := n \hat{\beta}^*_{\mathrm{QA}}(w)^T (s'W)^{-1} \hat{\beta}^*_{\mathrm{QA}}(w) \Rightarrow \chi^2_p, \tag{13}$$

where χ^2_p is the chi-square distribution with p degrees of freedom.

We call the test proposed in Theorem 2 the WAQ test. Employing a WAQ estimator for combining distributional information is not a new technique. Koenker (1984) proposed and studied the weighted average of multiple RQ estimates. Recently, Kim *et al.* (2014) and Zhao and Xiao (2014) constructed efficient estimation by optimally weighting multiple RQ estimates. Unlike the test statistic at a single quantile in (9), we can take into account multiple quantiles simultaneously by using $\hat{\beta}_{QA}(w)$, so that the test statistic in (13) guarantees a more robust hypothesis test, not sensitive to the choice of τ and the distribution of Y_i.

The selection of weights could impact the performance of the proposed WAQ test. Therefore, we propose some reasonable choices of weights.

Approach (i): Uniform weights. The simplest choice is the uniform weights

$$w^{\mathrm{U}}_r = \frac{1}{t}, \quad r = 1, \ldots, t.$$

Approach (ii): Efficient weights. By reducing the variance, we may construct a powerful test. Kim *et al.* (2014) find the optimal weights in the sense that s' in (12) converges to the inverse of the Fisher information of f_Y, which is the well-known optimal Cramér–Rao bound as the number of quantiles goes to infinity. They also introduce the explicit expression of the optimal weights when uniformly spaced quantiles are used. Let $q_r = f_Y(Q_Y(\tau_r)), r = 1, \ldots, t$ and $q_0 = q_{t+1} = 0$. With uniformly spaced quantiles,

$$w^{\mathrm{E}}_r = \frac{(2q_r - q_{r-1} - q_{r+1})\, q_r}{\sum_{r=1}^{t} (2q_r - q_{r-1} - q_{r+1})\, q_r}, \quad r = 1, \ldots, t.$$

Approach (iii): Sparsity function weights. The reciprocal of a density function evaluated at the quantile of interest has been termed the 'sparsity function'. To construct a more powerful test, we may assign more weight to the quantiles whose $Q_e(\tau)$ are larger. Then, when an ARCH effect exists, the resultant standardized difference of $\hat{\beta}^*_{\mathrm{QA}}(w)$ from 0 may become larger, which leads to a more powerful test. Recall $\beta(\tau) = \beta Q_e(\tau)$. Because the mode of most random variables is close to the middle of its range, most of the area under the density function of $e_i = |\varepsilon_i|$ is typically near 0 and the density function gradually decreases as e_i increases. Thus, we propose weights assigned with respect to the sparsity function, which puts more weight on large quantiles. For practical convenience, Y_i is used instead of e_i to compute the sparsity function weights.

$$w^{\mathrm{S}}_r = \frac{1}{c f_Y(Q_Y(\tau_r))}, \quad c = \sum_{r=1}^{t} \frac{1}{f_Y(Q_Y(\tau_r))}, \quad r = 1, \ldots, t.$$

Remark 1. The aggregate estimate in (10) is also constructed by averaging single QR estimators from $\mathrm{argmin}_b \sum_i \rho_\tau \left(X_i - Z_i^T b\right)$ with the original response X_i instead of $|X_i|$. Then, when the density function of the errors ε_i is symmetric about 0, it makes sense to assign the same weight to the quantiles reflected at 0, as with the proposed weights. Then, the corresponding parameter, $\beta_{QA}(w) = \beta \sum_{r=1}^{t} w_r Q_\varepsilon(\tau_r)$, is always 0 regardless of

DOI: 10.1111/jtsa.12089

the value of β with uniformly spaced quantiles. By contrast, formulating the RQ estimator in terms of the absolute value $Y_i = |X_i|$ avoids the problem caused by symmetric density functions because $Q_e(\tau_r)$ are non-negative.

Remark 2. The proposed WAQ test is readily applied to the quadratic ARCH model in (2) with $Y_i = X_i^2$ and $e_i = \varepsilon_i^2$, and similarly all the asymptotic properties and test statistics earlier can be investigated. However, extreme returns that the quadratic ARCH model often generates cause difficulties in estimating density functions evaluated at large quantiles (i.e. $\tau = 0.8, 0.9$). As a result, the WAQ would be oversized in the quadratic ARCH model with heavy-tailed distributions, and hence, to alleviate the issue, it requires a larger sample size.

3.2. Quantile Stacking Test

In this section, we introduce another hypothesis test for ARCH effects by combining information across multiple quantiles. Koenker and Zhao (1996) noted that heteroscedasticity tests based on the joint asymptotic normality of multiple RQ estimators $\hat{\beta}(\tau_1), \ldots, \hat{\beta}(\tau_t)$ might be constructed. Following their idea, we consider the $t(p+1)$ vectors

$$\hat{\beta}_{\text{QS}} = \left\{ \hat{\beta}(\tau_1)^T, \ldots, \hat{\beta}(\tau_t)^T \right\}^T$$

and $\beta_{\text{QS}} = \left\{ \beta(\tau_1)^T, \ldots, \beta(\tau_t)^T \right\}^T$. $\hat{\beta}_{\text{QS}}$ contains information about the parameters of interest across $\tau_1, \ldots, \tau_t$. Furno (2004) also adopts this stacking idea to construct the RQ test for conditional heteroscedasticity. Denote by $\otimes$ the Kronecker product.

Theorem 3. Suppose that Assumption 1 holds. Define $\hat{\beta}^*_{\text{QS}} = \left\{ \hat{\beta}^*(\tau_1)^T, \ldots, \hat{\beta}^*(\tau_t)^T \right\}^T$.

(i) Then, the following asymptotic normality holds

$$\sqrt{n} \left(\hat{\beta}_{\text{QS}} - \beta_{\text{QS}} \right) \Rightarrow N(\mathbf{0}, G),$$

where $G = H \otimes \left(V_1^{-1} V_0 V_1^{-1} \right)$ with $H = \left\{ \frac{\min(\tau_r, \tau_{r'}) - \tau_r \tau_{r'}}{f_e(Q_e(\tau_r)) f_e(Q_e(\tau_{r'}))} \right\}_{1 \le r, r' \le t}$.

(ii) Recall W in Corollary 2. Under the null hypothesis in (3), we have

$$T_{QS} := n \left(\hat{\beta}^*_{\text{QS}} \right)^T G'^{-1} \hat{\beta}^*_{\text{QS}} \Rightarrow \chi^2_{pt}, \tag{14}$$

where $G' = H' \otimes W$ with $H' = \left\{ \frac{\min(\tau_r, \tau_{r'}) - \tau_r \tau_{r'}}{f_Y(Q_Y(\tau_r)) f_Y(Q_Y(\tau_{r'}))} \right\}_{1 \le r, r' \le t}$ and χ^2_{pt} is the chi-square distribution with pt degrees of freedom.

We call the test proposed in Theorem 3 the quantile stacking (QS) test.

3.3. Testing Procedure

So as to implement the WAQ test and the QS test, we need to estimate the unknown quantities $f_Y(Q_Y(\tau))$ and V_0. For this task, we propose the procedure later.

DOI: 10.1111/jtsa.12089

1. Estimate $Q_Y(\tau)$ by the sample τth quantile of Y_i denoted by $\hat{Q}_Y(\tau)$. Then, estimate $f_Y(Q_Y(\tau))$ by the non-parametric kernel density estimator:

$$\hat{f}_Y\left(\hat{Q}_Y(\tau)\right) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{\hat{Q}_Y(\tau) - Y_i}{h_n}\right), \tag{15}$$

   where $K(\cdot)$ is a non-negative kernel function and $h_n > 0$ is a bandwidth satisfying $h_n \to 0$ and $nh_n \to \infty$.
2. Estimate $V_0 = E\left(Z_1 Z_1^T\right)$ by $n^{-1}\sum_{i=1}^{n} Z_i Z_i^T$.
3. Test ARCH effects with the test statistics in (13) and (14).

In (15), we use the standard Gaussian kernel and adopt the 'rule of thumb' by Silverman (1986) for bandwidth selection

$$h_n = 0.9\min\{sd(Y_i), IQR(Y_i)/1.34\}n^{-1/5},$$

where $sd(Y_i)$ and $IQR(Y_i)$ are the sample standard deviation and sample interquartile range of Y_i respectively.

4. NUMERICAL RESULTS

In this section, we compare the following test statistics:

$T(0.5)$,	the single QR-based test with $\tau = 0.5$;
T_{QS},	the QS test;
$T_{QA}(w, t)$,	the WAQ test with weights w and t uniformly spaced quantiles;
LM,	the LM test (Engle, 1982);
LBS,	the LBS test (Lee and King, 1993).

The LM test statistic is nR^2, where n is the sample size and R^2 is the multiple correlation coefficient of $|X_i|$ and $(|X_{i-1}|, \ldots, |X_{i-p}|)$. Note that the original LM test statistic for the quadratic ARCH model in (2) derived in the work of Engle (1982) employs the multiple correlation coefficient of X_i^2 and $\left(X_{i-1}^2, \ldots, X_{i-p}^2\right)$ to obtain R^2. The LBS test statistic that is asymptotically standard normal under the null hypothesis of no ARCH effects is

$$\frac{\left\{(n-p)\sum_{i=p+1}^{n}\left(X_i^2/\hat{\sigma}^2 - 1\right)\sum_{j=1}^{p} X_{i-j}^2\right\} / \left\{\sum_{i=p+1}^{n}\left(X_i^2/\hat{\sigma}^2 - 1\right)^2\right\}^{1/2}}{\left\{(n-p)\sum_{i=p+1}^{n}\left(\sum_{j=1}^{p} X_{i-j}^2\right)^2 - \left(\sum_{i=p+1}^{n}\sum_{j=1}^{p} X_{i-j}^2\right)^2\right\}^{1/2}},$$

where $\hat{\sigma}^2$ is the maximum likelihood estimator of σ^2 assuming that the errors ε_i are i.i.d. $N(0, \sigma^2)$.

For comparison, two models are considered:

$$\text{(Model I) ARCH(1)}: \quad X_i = (1 + \beta_1|X_{i-1}|)\varepsilon_i, \quad i = 1, \ldots, n;$$
$$\text{(Model II) ARCH(2)}: \quad X_i = (1 + \beta_2|X_{i-2}|)\varepsilon_i, \quad i = 1, \ldots, n.$$

Model II is used to examine how powerful the tests are in detecting a remote ARCH effect. For the error process $\{\varepsilon_i\}_{1 \le i \le n}$, we consider various distributions to investigate the performance of the tests: the standard normal; the centred chi-square distribution with 4 degrees of freedom to examine the impact of skewness; the Student's

DOI: 10.1111/jtsa.12089

Table I. Empirical sizes and powers of $T_{QA}(w,t)$ in percentages for Model I with five distributions: N(0,1); centred chi-square distribution with 4 degrees of freedom (χ_4); Student's t-distribution with 3 degrees of freedom (t_3); normal mixture (mix): $0.5N(0,1) + 0.5N(0,4)$; standard laplace distribution (laplace)

		$n = 50$			$n = 100$				
Error	β_1	$T_{QA}(5)$	$T_{QA}(9)$	$T_{QA}(19)$	$T_{QA}(5)$	$T_{QA}(9)$	$T_{QA}(19)$	$T_{QA}(w^E,9)$	$T_{QA}(w^S,9)$
N(0,1)	0	3.60	3.95	4.12	4.30	4.30	4.49	1.77	4.66
	0.1	6.11	6.49	7.34	9.19	10.02	10.64	4.40	11.22
	0.2	13.69	15.10	15.77	24.94	27.10	29.19	10.69	29.68
	0.3	26.30	28.84	30.22	48.91	52.72	55.80	21.75	57.09
χ_4	0	3.90	4.33	4.40	4.39	4.79	4.78	3.77	6.63
	0.1	30.09	31.12	28.70	53.63	54.37	54.07	34.78	50.28
	0.2	69.83	70.56	66.13	94.48	94.55	94.49	74.61	91.43
	0.3	92.50	92.80	84.34	99.58	99.64	99.64	93.23	98.42
t_3	0	4.61	4.86	4.29	4.47	4.94	5.77	3.80	5.50
	0.1	14.34	15.34	14.10	23.93	24.32	25.39	14.33	20.93
	0.2	31.38	32.68	28.91	51.75	53.25	54.18	30.15	45.81
	0.3	49.03	50.04	44.18	76.33	77.52	78.21	49.06	70.38
Laplace	0	3.79	3.93	4.12	3.59	4.35	4.27	1.62	4.45
	0.1	10.90	11.45	11.44	16.46	17.19	18.25	5.65	16.99
	0.2	23.01	24.79	24.53	40.98	42.77	44.88	15.76	40.46
	0.3	39.54	42.28	40.71	64.91	68.98	70.18	29.73	65.50
Mix	0	3.69	3.88	3.70	4.51	4.27	4.49	1.65	4.18
	0.1	23.25	25.25	24.59	37.84	41.15	44.51	14.94	38.22
	0.2	53.25	57.55	54.38	78.43	82.33	84.79	48.56	78.68
	0.3	75.23	78.84	69.14	95.02	96.57	97.23	75.95	94.01

The nominal size is 5%.

t-distribution with 3 degrees of freedom; the standard laplace distribution; and a normal mixture distribution to study the impact of kurtosis. The simulations are implemented with 50 and 100 observations with 10,000 replications and investigate the empirical size and power properties of the tests with the nominal level of 5%.

4.1. Choice of the Number of Quantiles and Weights

Before comparing the proposed WAQ test with the likelihood-based tests, we explore the influence of the number of quantiles and weights on the WAQ test. For notational simplicity, let $T_{QA}(t) = T_{QA}(w^{\mathrm{U}}, t)$. So as to examine the impact of the number of quantiles, Table I reports the empirical sizes and powers of $T_{QA}(5), T_{QA}(9)$ and $T_{QA}(19)$ for Model I. Although $T_{QA}(9)$ and $T_{QA}(19)$ deliver the best performance with 50 and 100 observations respectively, the simulation results suggest that the number of quantiles is not a crucial factor in the performance of the WAQ test if it is not too small. It is worth mentioning that when we implemented the WAQ test with 50 observations for heavy-tailed distributions, $T_{QA}(19)$ sometimes produces extremely large values of s' in (12), which results in numerical singularity since the estimate of $f_Y(Q_Y(0.95))$ is almost 0 on occasion. To circumvent this computational issue, nine uniformly space quantiles are adopted for the comparison with the other tests.

From the results earlier and other practices on QR, a rule-of-thumb choice for the number of quantiles is proposed. Denote by $\lfloor v \rfloor$ the integer part of v. We recommend $t = \lfloor n/5 \rfloor$ for $n < 100$ and $t = 19$ for $n \geq 100$. Then, the WAQ test can avoid computational issues mentioned in the preceding paragraph and still make use of enough distributional information. We do not suggest using too many of equally spaced quantiles for a couple of reasons. First, the RQ estimates at extremely large quantiles contain considerably larger variability than the others. As a result, the test could be oversized especially with heavy-tailed distributions. Second, when τ is very close to 0, we may lose power because $Q_e(\tau) \approx 0$.

In addition to $T_{QA}(t)$, Table I also displays the empirical sizes and powers of $T_{QA}(w^E, 9)$ and $T_{QA}(w^S, 9)$ with 100 observations to study the effect of the choice of weights. Unlike the number of quantiles, the choice of weights has a strong influence on the performance of the WAQ test. Among the three weights considered, the uniform weights deliver the best overall performance, so they are used to compare with the other tests.

DOI: 10.1111/jtsa.12089

Table II. Empirical sizes and powers of the five autoregressive conditional heteroscedasticity effects tests considered in percentages for Model I

		$n = 50$					$n = 100$				
Error	β_1	$T(0.5)$	T_{QS}	$T_{QA}(9)$	LM	LBS	$T(0.5)$	T_{QS}	$T_{QA}(9)$	LM	LBS
N(0,1)	0	2.42	*4.11*	*3.95*	*4.74*	*4.36*	2.98	2.71	*4.30*	*4.58*	*4.29*
	0.1	3.97	7.10	6.49	7.10	**10.96**	6.00	7.35	10.02	10.75	**17.63**
	0.2	8.11	14.14	15.10	15.12	**22.83**	13.75	18.03	27.10	29.43	**37.32**
	0.3	13.62	24.07	28.84	28.11	**35.44**	26.59	37.00	52.72	55.08	**58.55**
χ_4	0	2.29	8.98	*4.57*	*3.53*	*4.48*	2.84	8.05	*4.79*	*4.09*	*4.73*
	0.1	18.84	29.82	**31.12**	23.08	22.12	31.15	46.35	**54.37**	44.82	34.91
	0.2	46.18	63.85	**70.56**	61.28	51.33	74.35	90.12	**94.55**	89.86	73.90
	0.3	72.84	88.23	**92.80**	85.07	71.25	94.06	99.25	**99.64**	98.73	88.41
t_3	0	3.33	10.48	*4.92*	3.16	*3.54*	*3.58*	9.32	*4.94*	3.44	*4.08*
	0.1	10.90	22.72	**15.34**	8.68	9.96	17.56	28.80	**24.32**	16.38	14.60
	0.2	22.05	37.95	**32.68**	20.95	20.36	37.90	52.36	**53.25**	41.05	30.16
	0.3	34.19	52.19	**50.04**	37.94	29.84	57.70	73.50	**77.52**	65.63	47.51
Laplace	0	3.36	8.16	*4.04*	*3.79*	*4.36*	*3.60*	*5.81*	*4.35*	*4.30*	*4.86*
	0.1	7.23	17.11	11.45	8.95	**11.60**	10.87	**18.94**	17.19	15.65	17.28
	0.2	15.18	30.43	**24.79**	20.12	21.62	24.92	39.63	**42.77**	38.90	36.78
	0.3	24.54	44.74	**42.28**	34.04	31.25	43.72	62.54	**68.98**	63.04	52.96
Mix	0	2.77	10.51	*4.05*	*3.59*	*4.27*	2.95	7.72	*4.27*	*4.51*	*5.02*
	0.1	13.73	32.67	**25.25**	19.76	21.98	24.59	42.21	**41.15**	39.00	39.15
	0.2	36.11	60.38	**57.55**	45.82	35.45	63.32	79.51	**82.33**	78.26	63.13
	0.3	60.07	79.46	**78.84**	67.89	46.28	87.39	95.45	**96.57**	94.39	72.95

LM, Lagrange multiplier; LBS, locally most mean powerful based score. The empirical sizes close to the nominal level, those inside the 3.5–6.5% range, are in italics, and the largest empirical powers among them are reported in bold.

$T_{QA}(w^S, 9)$ produces comparable performance with $T_{QA}(9)$ for the non-normal errors considered and conveys slightly better performance with the standard normal errors. Therefore, the sparsity function weights are an alternative choice when the errors are symmetric and not heavy tailed. It is interesting to note that $T_{QA}(w^E, 9)$ has much smaller power than the other two tests. To reduce the variance, weights are assigned mostly to small quantiles having smaller variance, so that the resultant test statistics are relatively small and consequently less powerful. Based on the previous discussion, we recommend the uniform weights in practice because of its simplicity and powerful performance.

4.2. Performance in the Quadratic ARCH Model

The empirical sizes and powers of the five tests listed in Section 4 are presented in Tables II and III. In Table II, as mentioned by Lee and King (1993) and references therein, the LM test is slightly undersized with a small sample size especially when non-normal errors are considered. In contrast, the empirical sizes of the WAQ test and the LBS test are quite close to the nominal level regardless of the type of the error distributions considered. For the standard normal distribution, the LBS test is most powerful although the WAQ test and the LM test have comparable performance. The result is reasonable because the LBS test is built using the normality assumption. For the other distributions considered, the WAQ test is most powerful and delivers superior performance compared with the other methods, and the LM test outperforms the LBS test. The QS test is often significantly oversized, but its empirical powers are almost the same as those of the WAQ test. Furno (2004) reports the similar empirical size results for the QS test.

It is interesting to see that for a remote ARCH effect (Model II), the WAQ test is the only one having good empirical sizes with 50 observations and non-normal errors. The LBS test is even more undersized than the LM test. Furthermore, even for the standard normal distribution, the WAQ test performs almost equivalently to the tests developed under the normality assumption. Overall, the WAQ test detects a remote ARCH effect more sensitively than the other tests. In summary, we conclude that the WAQ test significantly outperforms the widely used likelihood-based tests in the sense of its accurate size, larger power and robust performance to error distributions.

DOI: 10.1111/jtsa.12089

Table III. Empirical sizes and powers of the five autoregressive conditional heteroscedasticity effects tests considered in percentages for Model II

		$n=50$					$n=100$				
Error	β_2	$T(0.5)$	T_{QS}	$T_{QA}(9)$	LM	LBS	$T(0.5)$	T_{QS}	$T_{QA}(9)$	LM	LBS
N(0,1)	0	1.45	2.34	3.12	*4.33*	2.83	2.04	1.48	*3.83*	*4.69*	*3.86*
	0.1	2.40	3.79	5.16	**5.72**	6.07	4.16	3.62	8.12	9.34	**11.05**
	0.2	4.64	7.01	10.27	**11.20**	12.05	9.67	10.24	21.48	**23.24**	22.74
	0.3	8.72	13.28	20.82	**21.73**	17.99	18.23	21.82	42.39	**44.66**	36.78
χ_4	0	1.47	9.11	*4.61*	3.47	2.81	2.21	8.75	*4.75*	*3.62*	3.45
	0.1	13.41	21.16	**23.31**	16.28	12.35	24.61	33.07	**45.36**	36.35	22.22
	0.2	38.55	47.71	**62.07**	51.25	30.04	67.25	78.05	**90.66**	83.56	52.47
	0.3	66.24	76.62	**89.00**	79.09	48.83	91.34	97.25	**99.52**	97.12	72.95
t_3	0	3.34	10.52	*5.04*	2.93	2.17	*4.45*	9.54	*5.28*	*3.87*	3.12
	0.1	9.50	19.10	**13.12**	6.89	5.21	14.10	22.55	**20.70**	12.68	9.70
	0.2	17.27	29.52	**26.06**	15.60	10.34	30.53	42.07	**46.29**	32.93	19.76
	0.3	29.28	43.04	**43.34**	28.53	17.94	50.02	61.94	**70.63**	55.84	30.49
Laplace	0	2.67	7.05	*3.64*	3.45	2.69	3.11	*5.12*	*4.03*	*4.30*	*3.67*
	0.1	5.36	13.24	**8.67**	6.24	6.24	8.32	12.73	**13.08**	11.28	11.08
	0.2	11.06	22.64	**20.16**	14.78	11.84	19.72	28.51	**34.72**	31.10	22.16
	0.3	19.29	34.54	**35.09**	27.11	18.02	35.88	48.63	**60.40**	55.13	33.91
Mix	0	2.72	9.73	*3.75*	*3.64*	2.87	2.85	7.39	*3.79*	*3.90*	*3.70*
	0.1	9.71	26.53	**19.55**	14.46	11.46	18.00	32.09	**34.09**	30.41	24.09
	0.2	29.61	51.53	**49.91**	37.93	21.57	56.28	70.55	**77.14**	71.18	42.56
	0.3	52.69	71.36	**73.91**	57.97	25.77	83.63	91.20	**94.86**	90.42	51.41

LM, Lagrange multiplier; LBS, locally most mean powerful based score. The empirical sizes close to the nominal level, those inside the 3.5–6.5% range, are in italics, and the largest empirical powers among them are reported in bold.

5. PROOFS

Write $x_n \asymp y_n$ if $x_n/y_n \to 1$, $x_n = O(y_n)$ if $\sup_n |x_n/y_n| < \infty$, and $x_n = o(y_n)$ if $x_n/y_n \to 0$.

Proof of Theorem 1

(i) Recall $Y_i = |X_i|$, $Z_i = (1, Y_{i-1}, \ldots, Y_{i-p})^T$, $\beta = (\beta_0, \ldots, \beta_p)^T$ and $\beta(\tau) = Q_e(\tau)\beta$. Let $\Delta = \sqrt{n}(b - \beta(\tau))$ and $\xi_i = e_i - Q_e(\tau)$. Then, we can write

$$Y_i - Z_i^T b = \sigma_i e_i - Z_i^T b = \sigma_i \xi_i - \frac{Z_i^T \Delta}{\sqrt{n}}. \tag{16}$$

Because $\hat{\beta}(\tau)$ minimizes the criterion function in (4), by (16) $\hat{\Delta} = \sqrt{n}\left(\hat{\beta}(\tau) - \beta(\tau)\right)$ minimizes the following re-parameterized function of Δ:

$$\hat{\Delta} = \text{argmin}_\Delta \mathcal{L}(\Delta), \quad \mathcal{L}(\Delta) = \sum_{i=1}^{n} \left\{\rho_\tau\left(\sigma_i \xi_i - \frac{Z_i^T \Delta}{\sqrt{n}}\right) - \rho_\tau(\sigma_i \xi_i)\right\}.$$

Let $\delta_i(\Delta) = Z_i^T \Delta/\sqrt{n}$. Employing Knight's identity

$$\rho_\tau(u - \delta) - \rho_\tau(u) = -\delta\{\tau - \mathbf{1}(u \le 0)\} + \int_0^\delta \{\mathbf{1}(u \le s) - \mathbf{1}(u \le 0)\} ds,$$

we can write

$$\mathcal{L}(\Delta) = -A_n \Delta + I_n, \tag{17}$$

  J. Time. Ser. Anal. **36:** 26–38 (2015)
DOI: 10.1111/jtsa.12089

where

$$A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i^T \{\tau - \mathbf{1}(\sigma_i \xi_i \leq 0)\} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i^T \{\tau - \mathbf{1}(e_i \leq Q_e(\tau))\},$$

$$I_n = \sum_{i=1}^{n} \zeta_i, \quad \zeta_i = \int_0^{\delta_i(\Delta)} \mathbf{1}\{\sigma_i (e_i - Q_e(\tau)) \leq s\} - \mathbf{1}\{\sigma_i (e_i - Q_e(\tau)) \leq 0\}\, ds.$$

First consider I_n. Recall $V_1 = \mathbb{E}\left(Z_1 Z_1^T/\sigma_1\right)$. Using the double expectation formula,

$$\begin{aligned} \mathbb{E}(I_n) = \mathbb{E}\{\mathbb{E}(I_n|Z_i)\} &= \mathbb{E}\left[\sum_{i=1}^{n} \int_0^{\delta_i(\Delta)} \left\{F_e\left(Q_e(\tau) + \frac{s}{\sigma_i}\right) - F_e(Q_e(\tau))\right\} ds\right] \\ &\asymp \mathbb{E}\left\{\sum_{i=1}^{n} \frac{\delta_i^2(\Delta)}{2\sigma_i} f_e(Q_e(\tau))\right\} = \frac{f_e(Q_e(\tau))}{2} \Delta^T V_1 \Delta. \end{aligned} \tag{18}$$

Let $\gamma_k = \mathrm{cov}(\zeta_i, \zeta_{i+k})$. Then, for $\delta > 2$, $\gamma_0 \leq \mathbb{E}\left(\zeta_1^2\right) \leq \left\{\mathbb{E}\left(|\zeta_1|^\delta\right)\right\}^{2/\delta}$, and by Proposition 2.5 by Fan and Yao (2005), we have the α-mixing inequality $|\gamma_k| \leq 8\alpha_k^{1-2/\delta} \left\{\mathbb{E}\left(|\zeta_1|^\delta\right)\right\}^{2/\delta}$. Therefore,

$$\begin{aligned} \mathrm{var}(I_n) = n\gamma_0 + 2\sum_{k=1}^{n-1}(n-k)\gamma_k &\leq n\left(\gamma_0 + 2\sum_{k=1}^{n-1}|\gamma_k|\right) \\ &\leq n\left\{\mathbb{E}\left(|\zeta_1|^\delta\right)\right\}^{2/\delta}\left(1 + 16\sum_{k=1}^{\infty}\alpha_k^{1-2/\delta}\right). \end{aligned} \tag{19}$$

Notice that $|\zeta_i| \leq |\delta_i(\Delta)|\mathbf{1}(|\xi_i| \leq |\delta_i(\Delta)|/\sigma_i)$ and $|\delta_i(\Delta)| = O(1 + \sum_{j=1}^{p} Y_{i-j})/\sqrt{n}$. Then, by the dominated convergence theorem, $\mathbb{E}\left(|\sqrt{n}\zeta_i|^\delta\right) \to 0$, so we can show $\mathrm{var}(I_n) \to 0$. Thus, by (17)–(19), we have

$$\mathcal{L}(\Delta) = \frac{f_e(Q_e(\tau))}{2}\Delta^T V_1 \Delta - A_n \Delta + o_p(1).$$

By the quadratic approximation and the convexity lemma (Pollard, 1991), the desired asymptotic Bahadur representation follows:

$$\hat{\Delta} = \frac{V_1^{-1}}{\sqrt{n} f_e(Q_e(\tau))} \sum_{i=1}^{n} Z_i \{\tau - \mathbf{1}(e_i < Q_e(\tau))\} + o_p(1).$$

DOI: 10.1111/jtsa.12089

(ii) Since $Z_i\{\tau - \mathbf{1}(e_i < Q_e(\tau))\}$ are martingale differences with respect to the filtration $\mathcal{F}_i$ generated by $\{\varepsilon_i, \varepsilon_{i-1}, \ldots\}$, the desired asymptotic normality in (6) follows from Brown's (1971) martingale central limit theorem.

□

Proof of Theorem 2

By (5), we have the following asymptotic Bahadur representation:

$$\sum_{r=1}^{t} w_r \left\{\hat{\beta}(\tau_r) - \beta(\tau_r)\right\} = \frac{V_1^{-1}}{n} \sum_{i=1}^{n} Z_i \sum_{r=1}^{t} \frac{w_r}{f_e(Q_e(\tau_r))} \{\tau_r - \mathbf{1}(e_i < Q_e(\tau_r))\} + o_p(1).$$

Then, (11) is proved by the same argument as in the proof of Theorem 1 (ii) and $\mathrm{cov}\{\tau_r - \mathbf{1}(e_i < Q_e(\tau_r)), \tau_s - \mathbf{1}(e_i < Q_e(\tau_s))\} = \min(\tau_r, \tau_s) - \tau_r \tau_s$. (13) follows from the delta method. □

Proof of Theorem 3

By the Cramér–Wold theorem and the argument in the proof of Theorems 1 and 2, we can easily show the desired result. We omit the details here. □

ACKNOWLEDGEMENT

We are grateful to the editor and three anonymous referees for their constructive comments.

REFERENCES

Bollerslev T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**: 307–327.

Brown BM. 1971. Martingale central limit theorem. *Annals of Mathematical Statistics* **1**: 59–66.

Dufour JM, Khalaf L, Bernard JT, Genest I. 2004. Simulation-based finite-sample tests for heteroskedasticity and ARCH effects. *Journal of Econometrics* **122**: 317–347.

Engle RF. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**: 987–1007.

Fan J, Yao Q. 2005. *Nonlinear Time Series: Nonparametric and Parametric Methods*. New York: Springer.

Furno M. 2004. ARCH tests and quantile regressions. *Journal of Statistical Computation & Simulation* **74**: 277–292.

Hong Y, Shehadeh RD. 1999. A new test for ARCH effects and its finite-sample performance. *Journal of Business & Economic Statistics* **17**: 91–108.

Kai B, Li R, Zou H. 2010. Local composite quantile regression smoothing: an efficient and safe alternative to local polynomial regression. *Journal of the Royal Statistical Society: Series B* **72**: 49–69.

Kim S, Zhao Z, Xiao Z. 2014. Efficient estimation for time-varying coefficient longitudinal models. *Manuscript*.

Koenker R. 1984. A note on L-estimators for linear models. *Statistics and Probability Letters* **2**: 323–325.

Koenker R, Zhao Q. 1996. Conditional quantile estimation and inference for ARCH models. *Econometric Theory* **12**: 793–813.

Lee JHH, King ML. 1993. A locally most mean powerful based score test for ARCH and GARCH regression disturbances. *Journal of Business & Economic Statistics* **11**: 17–27.

Pollard D. 1991. Asymptotic for least absolute deviation regression estimators. *Econometric Theory* **7**: 186–199.

Portnoy S, Koenker R. 1989. Adaptive L-estimation of linear models. *Annals of Statistics* **17**: 362–381.

Silverman BW. 1986. *Density Estimation*. London: Chapman and Hall.

Weiss AA. 1984. ARMA models with ARCH errors. *Journal of Time Series Analysis* **5**: 129–143.

Xiao Z, Koenker R. 2009. Conditional quantile estimation for generalized autoregressive conditional heteroscedasticity models. *Journal of the American Statistical Association* **104**: 1696–1712.

Zhao Z, Xiao Z. 2014. Efficient regressions via optimally combining quantile information. *Econometric Theory*. http://dx.doi.org/10.1017/S0266466614000176.

Zou H, Yuan M. 2008. Composite quantile regression and the oracle model selection theory. *Annals of Statistics* **36**: 1108–1126.

DOI: 10.1111/jtsa.12089