Journal of Mathematics and Statistics 10 (2): 169-185, 2014
ISSN: 1549-3644
© 2014 Science Publications
doi:10.3844/jmssp.2014.169.185 Published Online 10 (2) 2014 (http://www.thescipub.com/jmss.toc)

A NOVEL ROBUST AND EFFICIENT TOOL FOR DETECTING HETEROSCEDASTICITY

Wei Xiong and Maozai Tian

Center for Applied Statistics, School of Statistics Renmin University of China, Beijing, 100872, China

Received 2014-01-05; Revised 2014-02-22; Accepted 2014-03-21

ABSTRACT

One of the greatest values of Quantile Regression (QR) is that it provides a good procedure in the sense that QR could be much more efficient and sometimes arbitrarily more efficient in recovering the mean function than the Least Squares (LS) even when without moment conditions. However, heteroscedasticity definitely causes conditional variances of parametric or nonparametric estimates of mean functions to be large, sometimes this may lead to a great loss of efficiency of estimators and affect the goodness-of-fit test substantially and pratically conditional variance of data is of more concerned in statistical analysis these days, thus detecting heteroscedasticity before further analysis becomes essential. The virtue of QR as well as the limitation of LS motivates us to develop a new robust detecting tool for heteroscedasticity. Main contributions of this study include three aspects: First of all, a new Dynamic Quantile Regression (DQR) is introduced. Based on this method estimators for mean function, heteroscedastic function and the error distribution can be obtained simultaneously. Second, a novel diagnostic tool is developed for checking heteroscedasticity by employing the hybrid of QR and DQR. Theoretical properties of the procedure are investigated and we also demonstrate the performance of the new tool on small sample power properties. Third, further estimator of the conditional variance can be obtained based on improved DQR, when heteroscedasticity is detected. Finally these methods are illustrated with some simulated examples. Compared with the classical testing procedures, Monte Carlo simulations indicate that the new tool is more effective, powerful and easy to implement. Applications to a real data analysis is also discussed.

Keywords: Heteroscedasticity, Dynamic Quantile Regression, Inference Quantile Process, Conditional Variance, Nonparametric Volatility

1. INTRODUCTION

Consider a general linear regression model with heteroscedasticity Equation 1.1:

$$Y = X^T\beta + \sigma(X)\varepsilon \tag{1.1}$$

where, Y is the response variable, X is a covariate. The coefficient β is unknown and the conditional variance function $\sigma^2(x) = \mathrm{Var}(Y|X = x)$ is used to model the heteroscedasticity. The error term ε in (1.1) is assumed to be independent of X and have mean 0 and variance 1.

Linear regression model is used extensively in statistical applications. A standard assumption for it is the homogeneity of error variances and some papers even assume that errors are normally distributed. Violation of these assumptions may invalid many of the traditional statistical analysis techniques and can lead to inefficiency of estimators. Thus to detect heteroscedasticity of a linear model is of crucial importance. No surprise to see that many diagnostic tools and statistical testing methods exist in literature for this, (Anscombe, 1961; Atkinson, 1985; Bickel, 1978; Cook and Weisberg, 1983), but note that most of them are

Corresponding Author: Maozai Tian, Center for Applied Statistics, School of Statistics Renmin University of China, Beijing, 100872, China Tel: (86)10 82500173

established in mean regression framework. While, Koenker and Bassett Jr (1982) considered an alternative approach based on regression quantiles, which is more robust to outliers. Later Wicox and Keselman (2004) made an improvement on Koenker's method and made it perform well in small sample size. However, many of the techniques adopted in detecting heteroscedasticity can be difficult in implementation and time consuming. Thus one of the objectives of this study is to develop an efficient and powerful diagnostic tool which can have a wide application. In addition to make the new proposed test statistic more robust, quantile regression technique is also applied in our methods.

Our new method is built on the quantile regression estimator and Dynamic Quantile Regression (DQR) estimator which is used for simultaneously estimating the mean regression function, conditional variance function and error distribution in nonparametric regression model. It has been shown that the DQR estimator was much more efficient than the least squares estimator and performed well even in the worst case scenario, for example when errors follow cauchy distribution.

Furthermore, the DQR estimator is computationally faster and easier than the Composite Quantile Regression (CQR) estimator (Zou and Yuan, 2008). These nice theoretical properties of DQR estimators motivate us to construct a diagnostic tool based on it. In addition, to make it a safer and more effective testing tool, related inference process is developed. Asymptotic properties of test statistic and inference process are also investigated. To examine the feasibility and efficiency of the new diagnostic tool, power properties of different sample size are studied later.

In real analysis, we would like to not only detect but model the conditional variance of data. Thus two heteroscedastic models are considered here, σ(x) is linear and σ(x) is nonparametric. DQR tecniques then can be applied to the estimation. For the nonparametric case, local linear technique is employed in the estimation procedure and some improvements are also made based on those previouds methods. Specifically, tuning parameters are not required to be selected beforehand in this study, which largely improves the efficiency of estimators and simplifies the calculation. Moreover, distribution of error ε can be determined simultaneously.

The rest of paper is organized as follows. In section 2, we give the basic idea of DQR method and study its asymptotic properties. Based on the DQR method, diagnostic tools including related inference process are developed in section 3. Section 4 presents the further estimation procedure of conditional variance function and distribution of error of a linear model. Two heteroscedastic cases are considered, σ(x) is linear and σ(x) is nonparametric. Asymptotic properties of estimators are also studied. In section 5, Monte Carlo simulations are conducted to examine the performances of the new diagnostic tool and efficiency of the estimators. Real data analysis is also presented in the end to give an illustration.

2. MATERIALS AND METHODS

In DQR procedure, the quantile τ is supposed to be a random variable, uniformly distributed in (0, 1) rather than a fixed constant in quantile regression, that is τ~U(0, 1). Thus, the quantile τ is like a dynamic ball rolling back and forth in interval (0,1) and that is the origin of the name Dynamic Quantile Regression (DQR).

Suppose that $\{(X_i, Y_i), i = 1,...,n\}$ is an independent and identically distributed random sample coming from model (1.1), that is:

$$Y_i = X_i^T\beta + \sigma(X_i)\varepsilon i$$

The errors $\{\varepsilon i\}_{i=1}^n$ are assumed to be independent and identically distributed with an unknown distribution F and E(εi) = 0, Var(εi) = 1. Then with the assumption τ ~U(0,1) it is obvious that $c_\tau \square F^{-1}(\tau)$ is also a random variable and $c_\tau \underset{\square}{\mathrm{id}} F^{-1}(\tau)\eta \square F$ thus Equation 2.1:

$$E_\tau(c_\tau) = 0, Var_\tau(c_\tau) = 1 \tag{2.1}$$

Furthermore, the τ-th conditional quantile regression function of response Y_i:

$$Q_\tau(Y_i|X_i) = X_i^T\beta + \sigma(X)c_\tau$$

Is a random variable and we have:

$$E_\tau[Q_\tau(Y_i|X_i)] = X_i^T\beta$$

Now we give the dynamic quantile regression estimation procedure. For randomly sampled quantiles $\{\tau_k, k = 1, 2,...,N\}$ from uniform distribution, the τ_k-th conditional quantile regression function of the response Y_i is:

$$Q_{\tau k}(Y_i|X_i) = X_i^T\beta + \sigma(X_i)c_{\tau k}$$

where, $c_{\tau k} = {}^{F-1}(\tau k)$. Then employ linear quantile regression technique introduced by Koenker (2005), we have:

$$\hat\beta_{\tau k} = \operatorname{Arg}\min \sum_{i=1}^{n} \rho_{\tau k}(Y_i - X_i^T b), k = 1,2,\ldots,N$$

where, $\rho_{\tau k}(u) = \tau_k u I(u \geq 0) + (\tau_k\text{-}1)uI(u < 0)$ is the check function at τ_k-th quantile. Then we can obtain the DQR estimate of β, denoted as $\hat\beta_{DQR}$ Eqution 2.2:

$$\hat\beta_{DQR} = \frac{1}{N}\sum_{k=1}^{N}\hat\beta_{\tau k} \qquad (2.2)$$

- A1. The error distribution F has continuous density f, with f(u) uniformly bounded away from 0 and ∞
- A2. Let $X = (X_1,\ldots,X_n)$. There exist positive definite matrix D, such that:

$$\lim_{n\to\infty}\frac{1}{n}X^T X = D$$

- A3. Denote $\Gamma = \operatorname{diag}(\sigma(X_i))$ and elements $\sigma(X_i)$ are bounded away from 0 and ∞ and there exists a positive definite matrix G, such that:

$$\lim_{n\to\infty}\frac{1}{n}X^T\Gamma^{-1}X = G$$

Assumption A1-A3 are basically the same for establishing the asymptotic normality of a single quantile regression (Koenker, 2005) and establishing the asymptotic property of composite quantile regression (Zou and Yuan, 2008). Then under these conditions, we have the following results for the DQR estimator.

Theorem 1

Under regular conditions A1-A3, if N → ∞, then:

$$\sqrt{n}(\hat\beta_{DQR} - \phi)\underset{\rightarrow}{D} N(0, G^{-1}DG^{-1})$$

where, $\underset{\rightarrow}{D}$ denotes convergence in distribution

Remark 1

Several observations can be seen from Theorem 1. First whether a linear model is homoscedastic or heteroscedastic has little impact on estimator $\hat\beta_{DQR}$, because it is always an asymptotic unbiased estimator and note that this fact is important for the later construction of the test statistic. Second, estimator $\hat\beta_{DQR}$ is easier to derive and much more available than estimators like $\hat\beta_{CQR}$. Only simple linear quantile regression is needed in the calculation and the randomness of quantile τ bring additional convenience and interpretability. Third, compared with least square method, estimator $\hat\beta_{DQR}$ also has more gains especially under some heavy-tail distributions such as t distribution and cauchy distribution. To illustrate this we compute the asymptotic relative efficiency between LSE and DQR. Define $\operatorname{ARE}\left(\hat\beta_{DQR},\beta_{LS}\right) = \frac{\operatorname{MSE}\left(\hat\beta_{LS}\right)}{\operatorname{MSE}\left(\hat\beta_{DQR}\right)}$ and after straightforward calculations, we see that as the sample size n approaches ∞:

$$\operatorname{ARE}\left(\hat\beta_{DQR},\beta_{LS}\right) \to \sigma_f(N)^{-4/5} \qquad (2.3)$$

where, $\sigma_f(N) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{i=1}^{N}\frac{\tau_i\Lambda\tau_j - \tau_i\tau_j}{f(F^{-1}(\tau_i))f(F^{-1}(\tau_j))}$ N is the number of dynamic quantiles selected in the DQR procedure and F(·) and f(·) is the density function and cumulative distribution function of the error distribution respectively. From Equation (2.3) we can see that ARE depends only on the error distribution and the choice of dynamic quantile number N. Thus for some commonly seen error distributions, values of ARE can be directly derived, see **Table 1**.

Table 1. ARE $(\hat\beta_{DQR},\hat\beta_{LS})$ for some error distributions

	ARE $(\hat\beta_{DQR},\hat\beta_{LS})$				
Error distribution	N = 5	N = 10	N = 50	N = 100	N = 500
N(0, 1)	0.7971	0.9336	0.9272	0.9869	0.9981
Laplace	1.3243	1.2132	1.1236	1.0293	1.0082
t distribution with df = 5	1.1120	1.0406	1.0366	1.0212	1.0237
t distribution with df = 3	1.4294	1.2563	1.2314	1.2032	1.0927
0.95 N (0, 1)+ 0.05N(0, 10^2)	3.0258	3.1282	1.8513	1.3367	1.0429
0.90 N (0, 1)+ 0.10N(0, 10^2)	4.3725	4.0577	1.2939	1.1352	1.1122

Several things can be observed from **Table 1**. First, for normal distribution LSE is expected to have the best performance, while ARE $(\hat{\beta}_{DQR}, \hat{\beta}_{LS})$ is very close to 1 as N becomes larger. Second for all the non-normal distributions listed in **Table 1**, DQR estimator can have higher efficiency especially when N is small. Finally, when N is large such as N = 100, 500, all the ARE values are very close to 1, that is to say generally DQR estimator can have more gains compared with LSE estimator.

3. DETECTING TOOL

A standard assumption in regression analysis is the homogeneity of error variances. Whereas it is usually proved to be incorrect when confronted with reality. In this section, we would like to effectively detect heteroscedasticity of a linear model, then a new robust diagnostic tool is developed for it based on DQR method proposed before.

For model (1.1), the simple hypothesis is the homogeneity, that is:

$$H_0 : \sigma(x) = C$$

where, C is a constant. Without loss of generality, we can let C = 1. Thus under H_0, model (1.1) reduces to the iid case Equation 3.1:

$$Y_i = X_i^T \beta + \varepsilon_i, i = 1, 2, \ldots, n \tag{3.1}$$

Now in quantile regression framework, given a quantile τ_k, coefficient β in (3.1) can be estimated by solving (Zou and Yuan, 2008):

$$\left(\hat{c}_{\tau k}, \hat{\beta}_{\tau k}^{QR}\right) = \underset{c,\beta}{\operatorname{Arg\,min}} \sum_{i=1}^{n} \rho_{\tau k}(Y_i - c - X_i^T \beta)$$

And $\hat{\beta}_{\tau k}^{QR}$ is an unbiased estimator of β, for under mild conditions (Koenker, 2005):

$$\sqrt{n}(\hat{\beta}_{\tau k}^{QR} - \beta) \underset{\rightarrow}{D} N\left(0, \frac{\tau_k(1-\tau_k)}{f^2(c_{\tau k})} D^{-1}\right)$$

where, D is the positive definite matrix defined in condition A2.

In terms of the DQR estimate of β in model (3.1), $\hat{\beta}_{DQR}$, it is also unbiased and have a smaller variance compared with the QR estimator, due to the mechanism of DQR estimation procedure and we can obtain under H_0:

$$\sqrt{n}(\hat{\beta}_{DQR} - \beta) \underset{\rightarrow}{D} N(0, D^{-1})$$

Remark 2

The above result is a special case of Theorem 1. For the iid case (homoscedasticity), the diagonal matrix Γ in condition A3 turns into an identity matrix. Thus by a simple substitution, the above asymptotic normality can be established.

Now consider any fixed quantile τ_k, $0<\tau_k<1$, then estimator $\hat{\beta}_{\tau k}^{QR}$ is unbiased provided that the errors have common variances, however this does not necessarily hold when heteroscedasticity exists. But in either case, estimator $\hat{\beta}_{DOR}$ remains unbiased. Thus due to the nice properties of DQR estimators, tests of the hypothesis H_0 can be established based on the statistic:

$$T_n(\tau) = \frac{n}{\omega^2(\tau) + 2v(\tau) + 1} (\hat{\beta}_{\tau}^{QR} - \hat{\beta}_{DQR})^T D(\hat{\beta}_{\tau}^{QR} - \hat{\beta}_{DQR}) \tag{3.2}$$

where, τ ∈ (0, 1) is any fixed quantile, D is a positive definite matrix defined in condition A2, $\omega^2(\tau) = \tau(1-\tau)/f^2(F^{-1}(\tau))$, $v(\tau) = \int_0^\tau F^{-1}(t)dt / f(F^{-1}(\tau))$. Note that from Equation (3.2). if no heteroscedasticity exists in a model, statistic $T_n(\tau)$'s value could be very small for any quantile τ ∈ (0, 1), whereas this value could be considerably large given that the model is of some heteroscedasticity. However, it does not seem to be reasonable for just considering the value of $T_n(\tau)$ at a certain quantile τ_k to measure the discrepancy of two models.Thus to assess the discrepancy of two models more credibly, it is natural to consider a $T_n(\tau)$ process, which is a global measure over entire distribution and we present Theorem 2 to further investigate the performance of test statistic $T_n(\tau)$ and $T_n(\tau)$ process,.

Theorem 2

Under conditions A1-A3, for any fixed t ∈ [∈, 1-∈]:

$$T_n(t) \underset{\rightarrow}{D} X_p^2$$

where, ∈∈ (0, 1/2) and p is the dimension of β. Furthermore, for any index set I, I⊂(0, 1), consider a test process $\{T_n(\tau): \tau \in I\}$, then under the null hypothesis:

$$\sup_{\tau \in I} T_n(\tau) \underset{\rightarrow}{W} \sup_{\tau \in I} Q_p^2(\tau), \text{for } \tau \in I$$

where, $\underset{\rightarrow}{W}$ denotes weak convergence and $Q_p(t) = \|B_p(t)\| / \sqrt{t(1-t)}$ is generally referred to as a Bessel process of order p, $B_p(t) \sim N(0, t(1-t)I_p)$, $\|.\|$ denotes the normalized Euclidean norm.

Table 2. Critical Values for sup $Q_q^2(t)$

q	α = 0.01	α = 0.05	α = 0.10
1	13.01	9.84	8.19
2	16.44	12.93	11.20
4	21.54	17.56	15.62

Parameter p is the degrees of freedom, these critical values are used in Monte Carlo simulation of size and power tests in section 4

For any fixed t ∈ (0, 1), we have $Q_p^2(t) \square X_p^2$.

Remark 3

Critical values for sup $Q_p^2(t)$ have been presented by (De Long, 1981; Andrews, 1993) via simulations. In this study, we just list part of this in **Table 2** for later use of section 4.

Practically, the proposed $T_n(\tau)$ tests require estimation of $\omega^2(\tau)$, which is related with unknown error distribution. In this study, we suggest using a plug-in method, substituting the unknown distribution function with standard normal. That is, $\omega^2(t)$ can be estimated by $\hat{\omega}^2(t) = t(1-t)\phi^2(\Phi^{-1}(t))$; Likewise estimate of ν(t) can be obtained via $\hat{v}(t) = \frac{\Phi^{-1}(t/2)}{\phi(\Phi^{-1}(t))}t$.

4. FURTHER ESTIMATION

Once the heteroscedasticity of a model is detected, we would like to know how the conditional variance of the response Y varies with covariate X. In this section, we obtain a new efficient estimator of $\sigma^2(x)$ based on DQR method and we also assumed the two forms of σ(x), linear and nonparametric.

4.1. Linear form of σ(x)

Assume that the conditional variance Var(Y|X) has a linear association with covariate X. Then model (1.1) can be rewritten as:

$$Y = X^T\beta + (X^T\gamma)\varepsilon \tag{4.1}$$

This model has been considered by many statisticians (Koenker and Zhao, 1994; He, 1997; Koenker and Machado, 1999). In quantile regression framework, it is generally assumed that the distribution of error ε is known or the τ th quantile of ε is 0. In this part two cases are considered. First we would like to investigate the case when distribution of error ε is known, i.e., F is known; Then we relax the restriction, F is supposed to be unknown and obtain estimates of the conditional variance function as well as error distribution.

4.1.1. F is Known

If F is known, estimation procedure could be very simple and we propose two methods to estimate unknown coefficients β and γ. For τ ~ U(0, 1), the τ-th conditional quantile of response Y_i is Equation 4.2:

$$Q_\tau(Y_i \,|X_i) = X_i^T(\beta + \gamma c_\tau) \square X_i^T b(\tau) \tag{4.2}$$

where, $c_\tau \square F^{-1}(\tau)$

According to the mechanism of DQR estimation procedure, randomly sampled N quantiles from U(0, 1), denoted as {τ_i, i = 1,2...,N}, then for the given τ = τ_k, b(τ_k) can be obtained by employing linear quantile regression (Koenker, 2005) Equation 4.3:

$$\hat{b}(\tau_k) = \operatorname*{Arg\,min}_b \sum_{i=1}^n \rho_{\tau k}(Y_i - X_i^T b), k = 1,...N \tag{4.3}$$

where, $\rho_{\tau k}$ (·) is the check function at τ_kth quantile. Thus we can construct equations:

$$\begin{aligned} \hat{b}(\tau_1) &= \beta + \gamma c_{\tau 1} \\ \hat{b}(\tau_2) &= \beta + \gamma c_{\tau 2} \\ &\;.... \\ \hat{b}(\tau_N) &= \beta + \gamma c_{\tau N} \end{aligned} \tag{4.4}$$

With Equations (4.4), two estimation procedure can be established.

4.1.1.1. Direct Solution

We can see from Equation 4.4, γ can be solved by subtracting two adjacent equations, that is:

$$\hat{\gamma}_j = \frac{\hat{b}_{(\tau j+1)} - \hat{b}(\tau_j)}{c\tau_{j+1} - c\tau_j}, j = 1,..., N-1 \text{ and } \hat{\gamma}_N = \frac{\hat{b}(\tau_1) - \hat{b}(\tau_N)}{C_{\tau 1} - C_{\tau N}}$$

Then estimates of β, γ can be obtained by Equation 4.5 and 4.6:

$$\hat{\gamma} = \frac{1}{N}\sum_{j=1}^N \hat{\gamma}_j \tag{4.5}$$

$$\hat{\beta} = \frac{1}{N}\sum_{j=1}^N \left(\hat{b}(\tau_j) - \hat{\gamma} c_{\tau_j}\right) \tag{4.6}$$

4.1.1.2. Regression Analysis

Equations (4.3) are actually regression equations, thus to obtain the estimated results of β and γ, classical linear regression methods can be employed. To illustrate this, first we present some notations.

Denote $\gamma = (\gamma_1,...,\gamma_p)^T$, $\gamma_s = \gamma/(\gamma^T\gamma)$, the standardized version of γ, then cτ can be represented as $c_\tau = b^T(\tau)\gamma_s - \beta^T\gamma_s$.

Thus a simple linear regression model can be established Equation 4.7:

$$c_{\tau i} = \hat{b}^T(\tau_i)\gamma_s - \beta^T\gamma_s + e_i, i = 1,2,....N \qquad (4.7)$$

where, e_i is the error, assumed to be normally distributed with $N(0, \sigma^2)$ and σ is some constant, need not to be known. Equivalently, let $C_\tau = (c_{\tau 1},...,c_{\tau N})^T$, $B_\tau^T = (b(\tau_1),...,b(\tau N))$ is a p×N matrix, $B_0 = (\beta^T\gamma_s)$. $(1,1,...,1)^T$ is a N×1 vector and $e = (e_1,...,e_N)^T$, then (4.7) becomes:

$$C_\tau = \hat{B}_{\tau\gamma s} - B_0 + e$$

Then β and γ_s can be estimated by classical least square techniques:

Remark 4

It is regular to suppose that the error e is normally distributed, due to the key information of model having been extracted; and $\{c_{\tau i}\}_{i=1}^N$ in (4.4) is supposed to be known owing to the assumption of F is known.

4.1.2. F is Unknown

Generally, the error distribution F is unknown. In this situation, β, γ and even error density f_ϵ can be estimated based on DQR method. We would give in details and the basic idea of this estimation procedure is Equation 2.1.

For linear model (4.1), Equation 4.2 still holds when distribution F is unknown, while this time c_τ is unknown. Since τ ~ U(0, 1), then we have:

$$E_\tau[b(\tau)] = \beta, \mathrm{Var}_\tau[b(\tau)] = \gamma\gamma^T$$

Randomly sampled N quantiles from uniform distribution on (0, 1), denoted as $\{\tau_k$: k = 1,2,...,N}, $\{\hat{b}(\tau_k)$: k = 1,2,...,N} can be estimated via Equation 4.3. Let $\Sigma_B = \mathrm{Var}_\tau(b(\tau))$. Thus we can obtain estimates of β and Σ_B through Equation 4.8 and 4.9:

$$\hat{\beta}_{DQR} = \frac{1}{N}\sum_{k=1}^{N}\hat{b}(\tau_k) \qquad (4.8)$$

$$\hat{\Sigma}_{DQR} = \frac{1}{N}\sum_{k=1}^{N}(\hat{b}(\tau_k) - \hat{\beta}_{DQR})(\hat{b}(\tau_k) - \hat{\beta}_{DQR})^T \qquad (4.9)$$

And $\hat{\gamma}$ can be obtained via $\hat{\Sigma}_B$. Furthermore, we can get $\{\hat{c}_{\tau k} : k = 1,...,N\}$:

$$\hat{c}_{\tau k}^2 = \frac{(\hat{b}(\tau_k) - \hat{\beta})^T(\hat{b}(\tau_k) - \hat{\beta})}{\mathrm{tr}(\hat{\Sigma}_B)}$$

where, tr(·) denotes the trace of a matrix. As c_τ and ε are samely distributed, thus by kernel density estimators, with the sample $\{\hat{c}_{\tau k} : k = 1,...,N\}$, density f_ϵ can be estimated:

$$\hat{f}_\epsilon(u) = \frac{1}{Nh}\sum_{k=1}^{N}K\left(\frac{u - \hat{c}_{\tau k}}{h}\right)$$

where, K(·) is a kernel function and h is the smoothing bandwidth.

4.2. Nonparametric

For model (1.1), the conditional variance function $\sigma^2(x)$ is usually unknown, thus in this section we would like to develop a nonparametric DQR method to obtain the estimate of $\sigma^2(x)$. Throughout this study, local linear regression techniques (Fan, 1993; Fan and Gijbels, 1996; Yu and Jones, 1998) are employed for nonparametric function.

Note that τ ~ U(0, 1), then the τ th conditional quantile of model (1.1) is:

$$Q_\tau(Y_i | X_i) = X_i^T\beta + \sigma(X_i)c_\tau$$

And it is also a random variable. Moreover, let $r_i = Y_i - X_i^T\beta$, thus the τth conditional quantile of r_i^2 is:

$$Q_\tau(r_i^2 | X_i) = \sigma^2(X_i)c_\tau^2$$

Consider expectations of both the two random variables $Q_\tau(Y_i|X_i)$ and $Q_\tau(r_i^2 | X_i)$, then we have:

$$E_\tau[Q_\tau(Y_i | X_i)] = X_i^T\beta, E_\tau[Q_\tau(r_i^2 | X_i)] = \sigma^2(X_i)$$

In local linear regression, consider estimating the value of $\sigma^2(x)$ at x_0, $\sigma^2(x)$ can be approximated locally by a linear function $\sigma^2(x) \approx \sigma^2(x_0) + \sigma^2(x0)(x-x_0)$ in the neighborhood of x_0, where $\dot{\sigma}^2(x)$ is the derivative of $\sigma^2(x)$. Thus estimation procedure can be conducted by the following three steps.

Step I: Estimate β

Based on the idea of DQR, β can be estimated by Equation 4.10:

$$\hat{\beta}_{DQR} = \frac{1}{N}\sum_{k=1}^{N}\hat{\beta}_k \tag{4.10}$$

where, $\hat{\beta}_k$ $\text{Argmin}_\beta \sum_i \rho_{\tau k}\ (Y_i - X_i^T\beta), k = 1, 2..., N, \{\tau_k\}_{k=1}^N$ are N dynamic quantiles randomly sampled from U(0, 1).

Step II: Obtain $\hat{\sigma}^2(x)$ via local linear DQR

The residuals are $\hat{r}_i = Y_i - X_i^T\hat{\beta}_{DQR}, 1, 2..., n$. By local linear quantile regression, $\{(\sigma_k^2(x_0), \dot{\sigma}_k^2(x_0)) : k = 1, ..., N\}$ can be obtained by solving the following equation:

$$(\hat{\sigma}_k^2(x_0), \hat{\dot{\sigma}}_k^2(x_0))$$
$$= \underset{\alpha_1,\alpha_2}{\text{Arg}\min}\sum_{i=1}^{n}\rho_{\tau_k}(\hat{r}_i^2 - \alpha_1 - \alpha_2(X_i - x_0))K\left(\frac{X_i - x_0}{h_1}\right), k = 1, ...N$$

where, $K(\cdot)$ is the kernel function and h_1 is a smoothing bandwidth. Then we have equation 4.11:

$$\hat{\sigma}_{DQR}^2(x_0) = \frac{1}{N}\sum_{k=1}^{N}\hat{\sigma}_k^2(x_0), \hat{\dot{\sigma}}_{DOR}^2(x_0) = \frac{1}{N}\sum_{k=1}^{N}\sigma_k^{\hat{\cdot}2}(x_0) \tag{4.11}$$

Step III: Estimate error distribution f_ε

$\{\hat{c}_{\tau k} : k = 1, ..., N\}$ can be obtained by employing quantile regression Equation 4.12:

$$\hat{c}_{\mathbf{k}} = \underset{c}{\text{Arg}\min}\sum_{i=1}^{N}\rho_{\tau_k}(\hat{r}_i - \hat{\sigma}(X_i)c)\ k = 1, ..., N \tag{4.12}$$

Then use kernel density estimators, we can obtain $\hat{f}_\varepsilon$ as follows Equation 4.13:

$$\hat{f}_\varepsilon(u) = \frac{1}{Nh_2}\sum_{k=1}^{N}K\left(\frac{u - \hat{c}_{\tau k}}{h_2}\right) \tag{4.13}$$

where, h_2 is a bandwidth.

Remark 5

Compared with the DQR method, we have a little improvement on the estimation of $\{\hat{c}_{\tau k}\}_{k=1}^N$. Estimation can be proceeded directly via (4.12) without bandwidth selection and local approximation, which make it greatly simplified and in our estimation procedure only two bandwidths h_1 and h_2 need to be selected. There are many effective methods existing to select a kernel density estimator bandwidth h_2, such as plug-in method (Silverman), cross-validation (Hardle). Throughout this study, we apply the rule of thumb bandwidth, that is $h_2 = 1.06 \min\{\hat{\sigma}, R/1.34\}n^{-1/5}$ where σ is standard deviation and R is the interquantile range. To select bandwidth h_1, we use the automatic bandwidth selection considered in Xiong *et al*. (2012), for different quantiles $\{\tau_k\}_{k=1}^N$, the optimal bandwidth $h_{\tau_k}^{opt} = l(\tau_k)h_{LS}$ where l(p) = {2p(1-p)/(ϕ($(^{\Phi-1}(p))$ $^{\Phi-1}(p))^2\}^{1/5}$, $\phi(\cdot), \Phi\ (\cdot)$ denotes the density and cumulative distribution function respectively and h_{LS} can be selected by some sophisticated methods. For the choice of dynamic quantile number N in each estimation procedure, we choose N = 100 or N = 500 as suggested in Xiong *et al*. (2012).

4.3. Asymptotic Distribution

Denote $p(\cdot)$ as the marginal density of X. Let $f(\cdot)$ and $F(\cdot)$ be the density and cumulative distribution function of error ε. The kernel function $K(\cdot)$ is symmetric with a bounded support and denote:

$$\mu_2(K) = \int u^2K(u)du, R(K) = \int K^2(u)du$$

Then we have the following theorems.

Theorem 3

Suppose that σ(x) has a linear form, σ(x) = x^T γ, the error distribution F is unknown and suppose that the regular conditions are satisfied, then as N → ∞:

$$\sqrt{n}(\hat{\beta}_{DQR} - \beta)\underset{\rightarrow}{D}N(0, \Omega),$$
$$\sqrt{N}\{\hat{\Sigma}_{DQR} - \Sigma_B - \frac{1}{n}(\theta - 1)\Omega\}\underset{\rightarrow}{D}N(0, \Sigma_B^2\lambda^2)$$

Where:

$$\Omega = G^{-1}DG^{-1}, \theta = \int_{\square}F(t)(1 - F(t))dt, \lambda^2 = E(\varepsilon^2 - 1)^2, \varepsilon\ F \text{and}\ \varepsilon = (Y - X^T\beta)/X^T\lambda$$

Theorem 4

Suppose that σ(x) is unknown, x_0 is an interior point of support of $p(\cdot)$. Error distribution F is unknown, then under the regular conditions given in appendix, if $h_1 \to 0, N \to \infty$, $nh_1 \to \infty$, then:

$$\sqrt{n}(\hat{\beta}_{DQR} - \beta) \underset{\rightarrow}{D} N(0, G^{-1}DG^{-1}),$$

$$\sqrt{nh_1}\{\hat{\sigma}^2_{DQR}(x_0) - \sigma^2(x_0) - \frac{1}{2}\mu_2(K)\ddot{\sigma}^2(x_0)h_1^2\} \underset{\rightarrow}{D} N\left(0, \frac{R(K)\sigma^4(x_0)}{p(x_0)}\lambda^2\right)$$

where, $\ddot{\sigma}^2(x_0)$ denotes the second derivative of $\sigma^2(x), \lambda^2 = E(\varepsilon^2 - 1)^2, \varepsilon = (Y - X^T\beta) / \sigma(X)$.

5. ILLUSTRATIVE EXAMPLES

In this section we continue to explore the behavior of the DQR estimator introduced in the previous section. We first use Monte Carlo simulations to study the performance of test statistic $T_n(\tau)$ and $T_n(\tau)$ process under hypothesis H_0 and H_1 and evaluate size and power of $T_n(\tau)$ at some certain quantiles for several dynamic quantile numbers. Then to assess the finite sample performance of the estimation procedures proposed, two different heteroscedastic models are considered respectively in example 2 and example 3. Throughout this section, Gaussian kernel is applied, i.e., $K(u) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}u^2\right)$ and we adopt the bandwidth selection scheme specified in Remark 5.

5.1. Example 1-Performance of Test Statistic $T_n(\tau)$

In the first example, we would like to investigate the performance of $T_n(\tau)$ defined in section 3. We generate two data each come from model H_0 and H_1 Equation 5.1:

$$H_0 : Y = 3X + \varepsilon, \leftrightarrow H_1 : Y = 3X + (X + X^2)\varepsilon \qquad (5.1)$$

With $X \sim U(0, 1)$, $\varepsilon \sim N(0, 1)$ and the sample size n = 200. In order to examine and compare the behavior of $T_n(\tau)$ process, choose the dynamic quantile number N = 10, 100, 500 respectively. First we would like to investigate the performance of the test statistic $T_n(\tau)$ at certain quantiles for different dynamic quantile numbers N, we calculate the type one error rate and power based on the hypothesis H_0 and H_1 at τ = 0.1, 0.25, 0.5, 0.75, 0.9. To obtain more accurate results, 500 simulations are conducted. Results are reported in **Table 3**.

From **Table 3**, it is clear that the sizes of the tests vary with dynamic quantile number N. For small N, such as N = 10, the tests are oversized; For large N, N = 500, the tests are almost undersized; and for moderate N, N = 100. These tests seem to perform better and it seems that to perform a satisfied test which can control the type one error rate, we can establish certain relationships between sample size n and dynamic quantile number N. We would investigate it further in the following. In addition, in terms of the power of the tests, something interesting would be found. For τ = 0.5, the power is almost 0 in different situations. Thus the alternative hypothesis is more difficult to discern. For all $\tau \neq 0.5$, the power is almost 1 under each cases especially when N is large, thus we can say the test tool is effective and of a success at $\tau \neq 0.5$. In fact it is accessible that the behavior of $T_n(\tau)$ at τ = 0.5 is poor, because under H_1 the linear conditional quantile function is $Q_{1/2}(Y|X) = X^T \beta$ due to $Q_{1/2}(\varepsilon) = 0$, thus the heteroscedastic function has no impact on $Q_{1/2}(Y|X)$ and test statistic $T_n(\tau)$ would behave similar under both hypotheses. Consequently, for some practical analysis if the prescribed quantile τ_k is not appropriate, then tests based on Tn(τk) may be difficult to distinguish. Then we suggest a more effective, safe and robust diagnostic tool $T_n(\tau)$ process, which measures the discrepancy over the entire distribution of τ.

In order to explore the effects sample size n had made on the tests, we consider sample size n = 50, n = 500 for model (5.1) respectively. Specially, we would like to give a rule to choose appropriate N in real analysis based on this.

Several features of **Table 4 and 5** merit attention. First, for small sample size n = 50, the sizes of the tests are smaller, while the power is very poor, that is the alternative is obviously more difficult to discern with small sample size; Second, the tests are a clear success at n = 500, where they have power near 1 in virtually all cases; Third, with larger N, the tests perform equally good under different sample sizes. Especially, compared with **Table 4** and **5** something interesting could be found. If we want to perform a good test which can control the type I error rate and also obtain a relatively large power, then N should be chosen of equal size with sample size n. For this, we consider a deviance $D_i = \left|T_n^{i+1}(\tau) - T_n^i(\tau)\right|, i = 1, 2, \ldots,$ where i represents the number of dynamic quantile employed in $T_n(\tau)$process. If D_i is sufficiently small, such as $D_i < \in$, for $\in \in (0, 1)$. Then we choose N = i for the test $T_n(\tau)$. This criterion also supports our inference, that is, in real analysis N and n could be of equal size.

To investigate the behavior of $T_n(\tau)$ process under model (5.1) with n = 200, plots are depicted in **Fig. 1**.

The $T_n(\tau)$ processes depicted in **Fig. 1** indicate a highly significant departure from the null hypothesis over the entire range of τ, whatever dynamic quantile number N is. When N is large, the results are more stable. For model H_0, we would expect that the $T_n(\tau)$ process to be nearly 0 over the entire range of $\tau \in [\in, 1-\in]$ and behave like the square of a normalized Brownian bridge process, this expectation is borne out for large dynamic quantile number N. Likewise we would like to find the $T_n(\tau)$ process under H_1 model is significantly different from 0 and it is also consistent with the figure plotted. Then we can say that the $T_n(\tau)$ process test statistic is a more diagnostic tool.

Moreover behaviors of test statistic $T_n(\tau)$ under other non-normal distributions may be of more concerned. Thus we would like to present the performance of $T_n(\tau)$ under model H_0 when ε follows Laplace and t_3 distribution respectively and results are plotted in **Fig. 2**.

The $T_n(\tau)$ processes in **Fig. 2** indicate a highly departure from null hypothesis over the entire range of τ for both non-normal distributions. Besides, with the increase of N, the $T_n(\tau)$ processes are more stable under both hypotheses.

5.2. Example 2-Linear Model, Under Different Errors

It is of interest to examine the property of our proposed DQR methods under different errors. In this example, we consider a totally linear case, that is σ(x) has a linear form. For this we generated 400 data sets, each consisting of n = 200 observations, coming from:

$$Y = 2X + (0.3X)\varepsilon$$

Table 3. Sizes and power of $T_n(\tau)$ tests, n = 200

		Type I error rate			Power size		
		0.1	0.05	0.01	0.1	0.05	0.01
τ = 0.1	N = 10	0.294	0.266	0.248	0.950	0.940	0.902
	N = 100	0.106	0.068	0.025	1.000	0.998	1.000
	N = 500	0.058	0.044	0.008	1.000	1.000	1.000
τ = 0.25	N = 10	0.336	0.202	0.031	0.740	0.690	0.646
	N = 100	0.102	0.062	0.024	0.956	0.900	0.812
	N = 500	0.044	0.026	0.004	0.994	0.958	0.850
τ = 0.5	N = 10	0.292	0.234	0.033	0.200	0.156	0.086
	N = 100	0.096	0.048	0.014	0.002	0.000	0.000
	N = 500	0.050	0.024	0.006	0.000	0.000	0.000
τ = 0.75	N = 10	0.318	0.224	0.033	0.738	0.712	0.640
	N = 100	0.098	0.062	0.022	0.952	0.930	0.800
	N = 500	0.054	0.026	0.006	0.996	0.978	0.838
τ = 0.9	N = 10	0.330	0.316	0.198	0.952	0.930	0.914
	N = 100	0.104	0.066	0.016	1.000	1.000	0.998
	N = 500	0.046	0.024	0.015	1.000	1.000	1.000

The table is based on 500 replications per cell. For the size and power, each cell reports the proportion of rejections of the designated test at the designated level of significance

Table 4. Sizes and power of $T_n(\tau)$ tests, n = 50

		Type I error rate			Power size		
		0.1	0.05	0.01	0.1	0.05	0.01
τ = 0.1	N = 10	0.144	0.138	0.094	0.548	0.478	0.336
	N = 100	0.070	0.038	0.012	0.630	0.492	0.274
	N = 500	0.052	0.040	0.006	0.660	0.470	0.252
τ = 0.25	N = 10	0.156	0.122	0.076	0.258	0.192	0.136
	N = 100	0.052	0.034	0.008	0.344	0.078	0.044
	N = 500	0.052	0.032	0.008	0.333	0.088	0.026
τ = 0.5	N = 10	0.128	0.108	0.080	0.010	0.010	0.010
	N = 100	0.062	0.024	0.006	0.000	0.000	0.000
	N = 500	0.032	0.020	0.006	0.000	0.000	0.000
τ = 0.75	N = 10	0.128	0.118	0.066	0.258	0.212	0.126
	N = 100	0.054	0.040	0.010	0.315	0.096	0.036
	N = 500	0.032	0.016	0.010	0.342	0.088	0.018
τ = 0.9	N = 10	0.136	0.038	0.080	0.562	0.494	0.334
	N = 100	0.054	0.040	0.014	0.638	0.468	0.030
	N = 500	0.054	0.022	0.012	0.616	0.508	0.268

Table 5. Sizes and power of $T_n(\tau)$ tests, n = 500

		Type I error rate			Power size		
		0.1	0.05	0.01	0.1	0.05	0.01
τ = 0.1	N = 10	0.482	0.434	0.378	0.986	0.988	0.970
	N = 100	0.180	0.126	0.090	1.000	1.000	1.000
	N = 500	0.096	0.040	0.008	1.000	1.000	1.000
τ = 0.25	N = 10	0.524	0.428	0.392	0.872	0.870	0.814
	N = 100	0.172	0.118	0.080	1.000	1.000	0.998
	N = 500	0.101	0.042	0.011	1.000	1.000	1.000
τ = 0.5	N = 10	0.500	0.452	0.410	0.432	0.314	0.268
	N = 100	0.190	0.108	0.074	0.008	0.006	0.000
	N = 500	0.074	0.048	0.086	0.000	0.000	0.000
τ = 0.75	N = 10	0.464	0.452	0.382	0.878	0.890	0.864
	N = 100	0.168	0.126	0.082	0.998	0.998	1.000
	N = 500	0.066	0.051	0.096	1.000	1.000	1.000
τ = 0.9	N = 10	0.466	0.402	0.332	0.992	0.986	0.970
	N = 100	0.182	0.114	0.068	1.000	1.000	1.000
	N = 500	0.074	0.050	0.011	1.000	1.000	1.000

The table is based on 500 replications per cell. For the size and power, each cell reports the proportion of rejections of the designated test at the designated level of significance

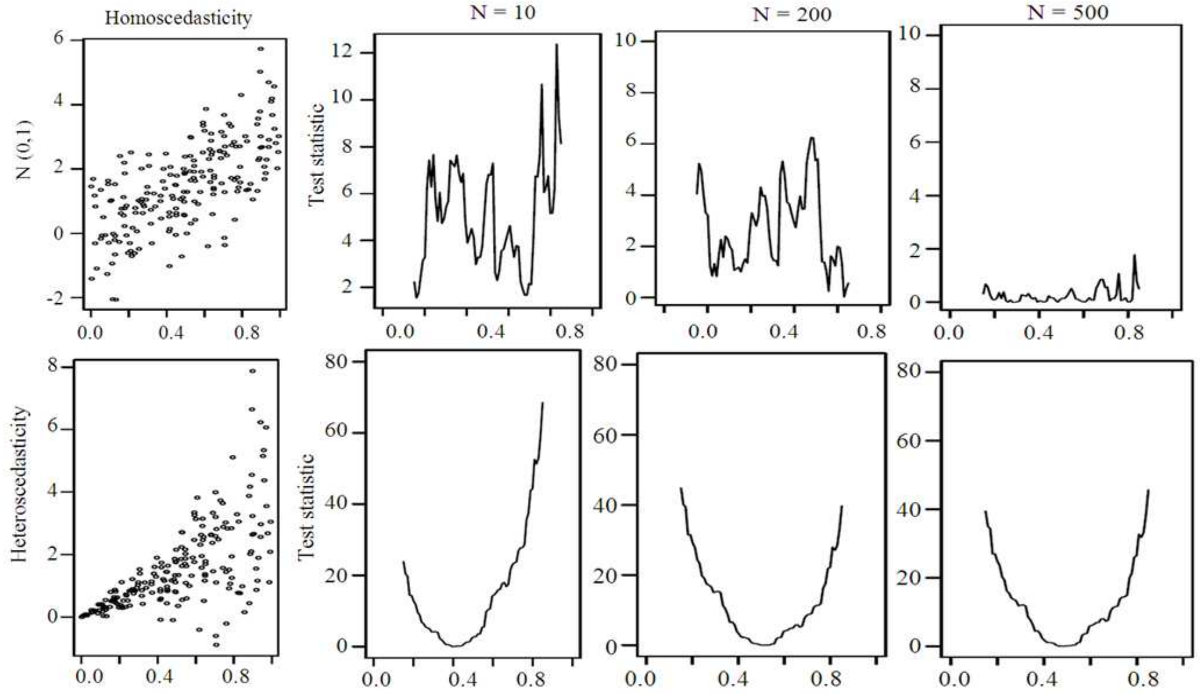

Fig. 1. Behavior of test statistic $T_n(\tau)$ under hypothesis H_0 and H_1 by choosing different dynamic quantile number N = 10, 100, 500

where, X ~ U(0, 1) and we considered five different error distributions for ε: N(0, 1), Laplace, t_3, t_5-distribution, Cauchy distribution. In the simulation, ε are scaled to have mean 0 and variance 1. For DQR estimator of β and γ, we consider dynamic quantile number N = 100, 500 respectively, as suggested in Xiong *et al*. (2012).

The mean and standard deviation of both $\hat{\beta}_{DQR}$ and $\hat{\gamma}_{DQR}$ over 400 simulations are summarized in **Table 6**.

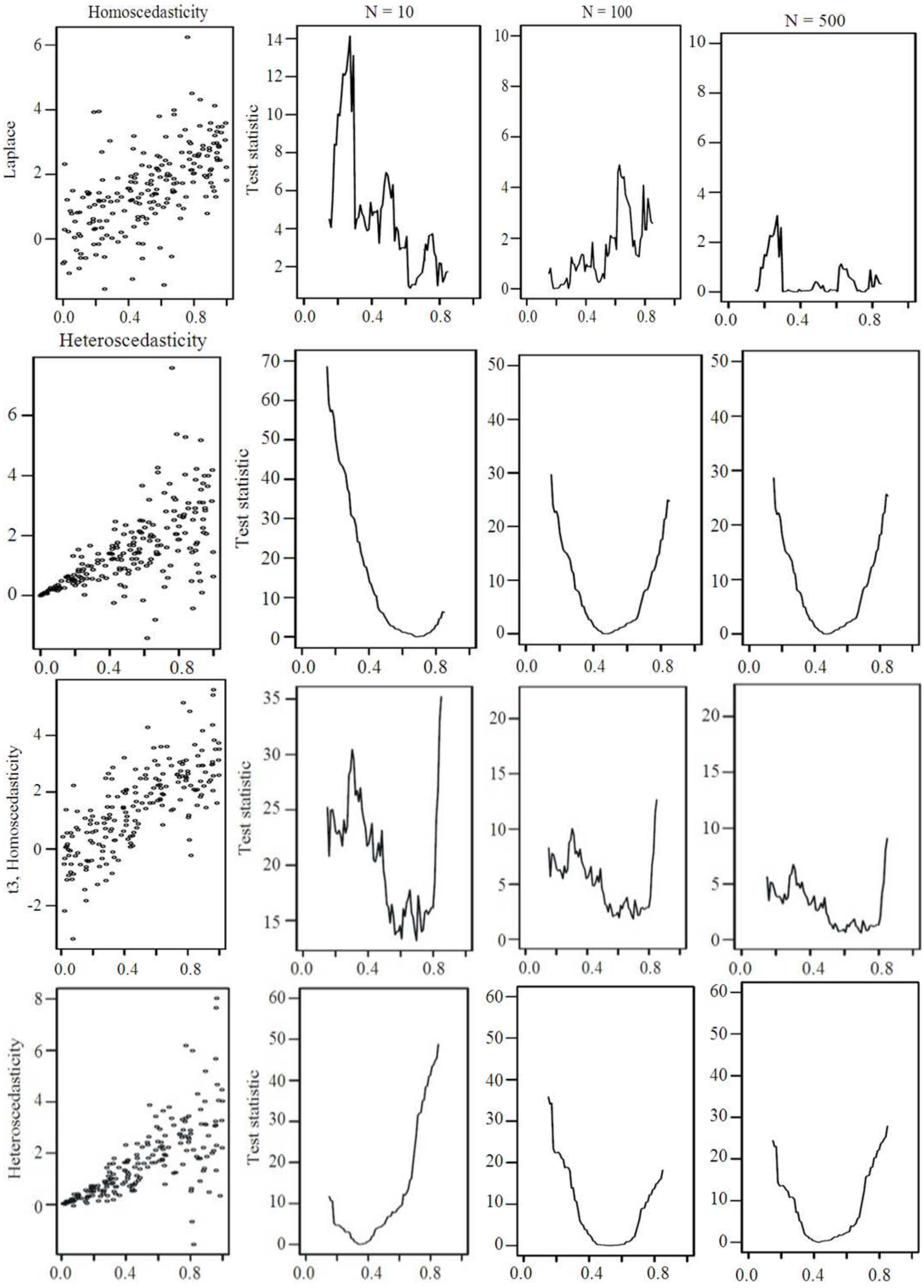

Fig. 2. Behavior of test statistic $T_n(\tau)$ under hypothesis H_0 and H_1 when ε ~ laplace and t3 respectively. The first two rows depicted the performances of $T_n(\tau)$ when error follows laplace distribution; the last two rows expressed the results under ε ~ t3 and the first column presents the scatter plots for each case, the last three columns each represents N = 10, 100, 500 respectively

Table 6. DQR estimates of coefficient β and γ

		β (= 2)		γ (= 0.3)	
Error distribution		Mean	Sd	Mean	Sd
N(0, 1)	N = 100	2.0004	0.0327	0.2992	0.0229
	N = 500	2.0015	0.0312	0.2988	0.0223
Laplace	N = 100	1.9979	0.0313	0.2968	0.0332
	N = 500	2.0000	0.0178	0.3000	0.0212
t_5	N = 100	2.0001	0.0311	0.2924	0.0360
	N = 500	2.0000	0.0180	0.2982	0.0209
t_3	N = 100	1.9999	0.0308	0.2910	0.0593
	N = 500	2.0006	0.0177	0.2985	0.0316
Cauchy	N = 100	2.0041	0.0334	0.2589	0.1693
	N = 500	2.0002	0.0176	0.2793	0.0917

Table 6 indicates that DQR estimator can well explore the heteroscedasticity of a model. Whatever the error distribution is, normal or non-normal, the estimated results is very close to the real value and with the quantile number N increases, this inference is obvious. Thus DQR estimator is robust. In addition, it is remarkably mentioning that even if the error follows a Cauchy distribution, DQR estimator can still capture the heteroscedasticity and obtain a good estimate, while it totally breaks down under the Least-square regression framework.

5.3. Example 3-Nonparametric Model

Generally, the conditional variance function $\sigma^2(x)$ is unknown. In this situation, local linear DQR method can be employed to obtain estimates of heteroscedastic function. To investigate the performance of local linear DQR estimator under nonparametric form, 400 simulation data sets are generated in this example and each consisting of n = 200 observations, coming from model:

$$Y = 3X + \frac{1}{4}(2 + \sin(2\pi X)) \in$$

where, X follows U(0, 1), $\sigma(x) = (x + \sin(2\pi x))/4$ and we choose $\in$ follow four different distribution: N(0, 1), Cauchy, t_3, t_5 distribution. In this example, we estimate coefficient β and $\sigma^2(x)$ over [0, 1]. The mean and standard deviation of ˆ $\hat{\beta}_{DQR}$ over 400 simulations are listed in **Table 7**, in which we also present the biases and standard deviation of $\hat{\sigma}^2_{DQR}(x)$ at x = 0.4. Likewise we choose dynamic quantile number N = 100 and 500 for comparisons. The estimated variance function for N = 100, 500 under three different error distributions: N(0, 1), Laplace and t_3 are depicted in **Fig. 3**, in which we also present the estimated error density by choosing N = 500.

Table 7 gives a good illustration of the capacity of DQR method to recover both coefficient β and variance function $\sigma^2(x)$, due to the small biases and standard deviations under different error distributions.

5.4. Real Data Analysis

For the real data, first we would like to detect whether there is certain variability of the data. For this, $T_n(\tau)$ process which proposed in the previous section is applied. If heteroscedasticity indeed exists, a heteroscedastic model could be built to model this.

In this section, we consider the Engel dataset employed by Koenker and Bassett Jr (1982) to test the heteroscedasticity. This dataset consists of 235 observations on household income and food expenditure for Belgian working class households. We use $T_n(\tau)$ process to detect whether there is variation in the data. Results are showed in **Fig. 4**.

Fig. 4 indicates a significant heteroscedasticity of the data. Then a linear model could be built to model the variation of data.

$$Y = X^T\beta + \sigma(X) \in$$

For the estimate of unknown coefficient β, we adopt two methods to make a comparison, classical least square method and newly dynamic quantile regression method. We choose N = 500 in DQR. Then we could obtain $\hat{\beta}_{DQR} = 0.630$, $\hat{\beta}_{LS} = 0.485$. The fitted linear regression model are plotted in **Fig. 5**.

From **Fig. 5**, it is evident that there exists an outlier in the data and the fitted linear model obtained by LS is highly depend on the outlier; whereas the result of DQR is more robust, it is of little impact on it.

Table 7. DQR estimates of coefficient β and $\sigma^2(x)$ at x = 0.4

		β(= 3)		$\sigma^2(x)$, x = 0.4		$\sigma^2(x)$
Error distribution		Mean	Sd	Bias	Sd	ASE
N(0, 1)	N = 100	3.0131	0.1964	0.0008	0.1232	0.0098
	N = 500	3.0048	0.0938	-0.0127	0.0900	0.0072
Laplace	N = 100	3.0014	0.1893	-0.0257	01440	0.0334
	N = 500	2.9986	0.0866	-0.0245	0.1526	0.0132
t_5	N = 100	3.0206	0.2272	-0.0090	0.1787	0.0258
	N = 500	2.9967	0.1066	-0.0194	0.1312	0.0176
t_3	N = 100	3.0076	0.1766	-0.0243	0.2043	0.0495
	N = 500	3.0023	0.0917	-0.0148	0.3200	0.0447

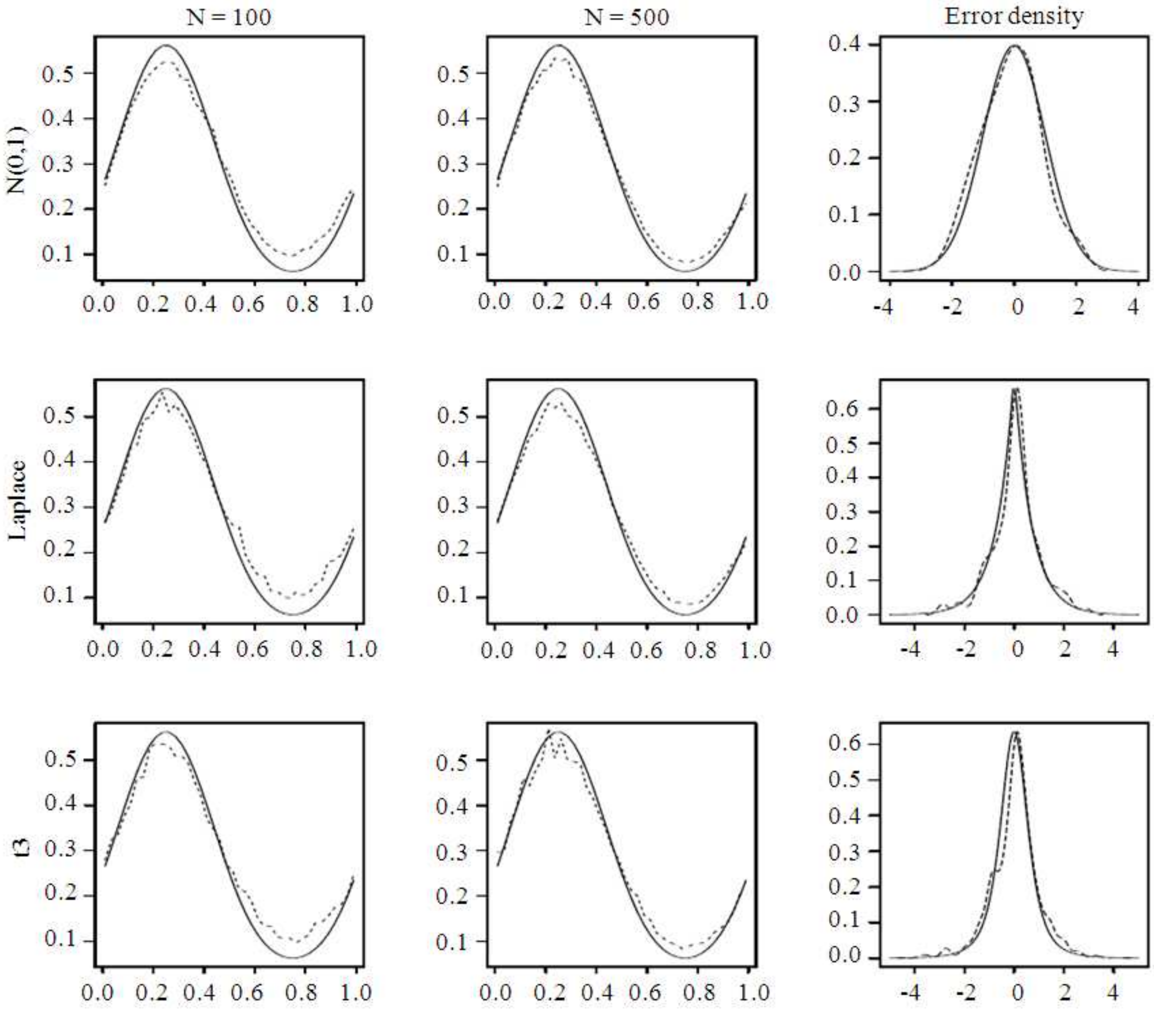

Fig. 3. Estimated results via local linear DQR method of Example 3. Three cases of error distribution are displayed in the figure. Each row represents the normal distribution, laplace and t_3 separately and dynamic quantile number N = 100, 500 for estimating the variance function $\sigma^2(x)$ are applied; For kernel error density estimators N is chosen to be N = 500; dash line --- denotes the estimated value and solid line -represents the real

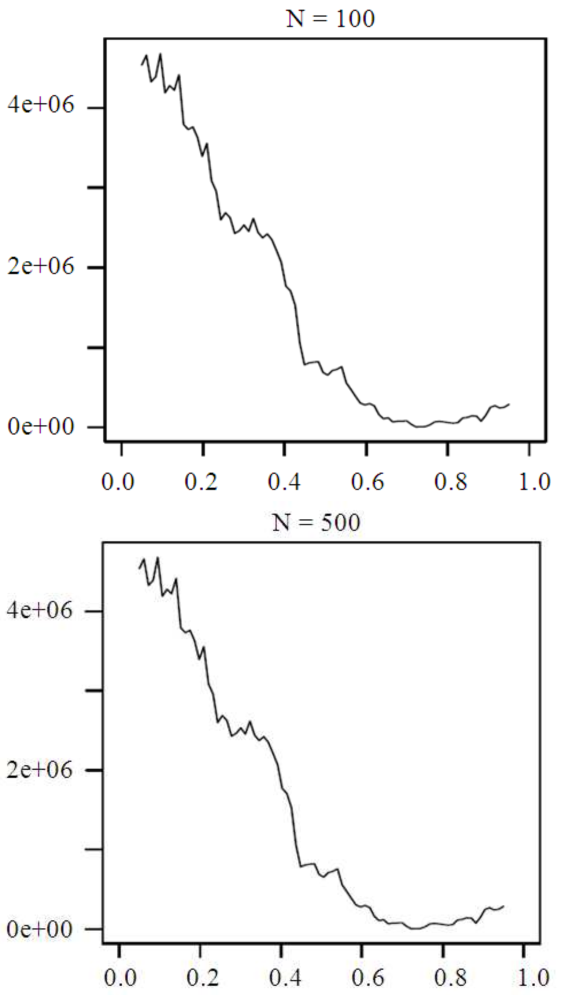

Fig. 4. Results of $T_n(\tau\ )$ for Engel Data. The dynamic quantile number N = 100 and 500 are chosen to make a comparison

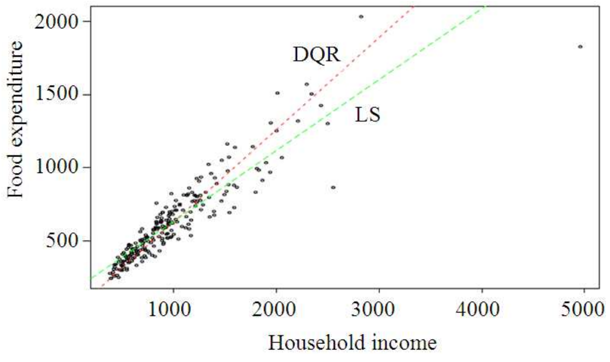

Fig. 5. Fitting a linear model via DQR and LS

6. CONCLUSION

In this study, we mainly consider the heteroscedastic model. To make a better data analysis, we first propose a robust method-Dynamic Quantile Regression (DQR) to give an efficient esimation. Then we develop a diagnostic tool which can effectively detect the heteroscedasticity of the datasets based on the hybrid of QR and DQR. To model the heteroscedastic function, two cases are considered, linear form and nonparametric form and for each we present the detailed estimation procedures and establish the asymptotic properties. Extensive Monte Carlo simulations are conducted to examine the finite performance of the proposed procedures. The results show that under various error distributions, DQR estimators outperform LS estimators and the Tn(τ) process could be a good alternative when detecting the heteroscedasticity. In addition, plots of Tn(τ) process give us a clear and direct awareness of the behavior of this statistic at different quantile points. The size and power of the test statisitc under different sample sizes also demonstrate the efficiency of the proposed methods. In empirical analysis, we apply the proposed DQR method as well as the Tn(τ) process to analyze the Engel dataset and we find that our methods could both effectively examine the heteroscedasticity and efficiently estimate the model compared with LS method.

Actually, the research can be extended to more general models with high dimensional covarites—the current hot issues. Furthermore, to avoid the so-called “curse of dimensionality”, we can apply the proposed method to semiparametric models which have more flexibility and interpretation. Further to explore the hidden structure and involve the dynamic feature, varying coefficient model can also be considered.

7. ACKNOWLEDGEMENT

The work was partially supported by National Natural Science Foundation of China (No.11271368), Beijing Philosophy and Social Science Foundation Grant (No.12JGB051), Project of Ministry of Education supported by the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20130004110007), The Key Program of National Philosophy and Social Science Foundation Grant (No. 13AZD064), Fundamental Research Funds for the Central Universities and the Research Funds of Renmin University of China (No.10XNK025) and China Statistical Research Project (No. 2011LZ031).

8. APPENDIX

Proof of Theorem 1 is included in proof of Theorem 3. See proof of Theorem 1, please refer to proof of Theorem 3 given in below.

Lemma 1.

For any fixed quantile $\tau_k \in (0, 1)$, let $\omega_{ij} = (\tau_i \Lambda \tau_j - \tau_i\tau_j)/(f(F^1(\tau_i))f(F^{-1}(\tau j)))$, $\tau \sim U(0, 1)$, then as $N \to \infty$:

$$\frac{1}{N}\sum_{j=1}^{N}\omega_{kj} \underset{\rightarrow}{a.s} v(k)$$

where, $\underset{\rightarrow}{a.s}$ Denotes convergence almost surely, $v_k = -\int_0^{\tau_k} F^{-1}(t)dt / f(F^{-1}(\tau_k))$ and F is a distribution function such that, for any random variable $X \sim F$, $E(X) = 0$ and $Var(X) = 1$.

Proof

As $\tau \sim U(0, 1)$ and $N \to \infty$, then $\frac{1}{N}\sum_{j=1}^{N}\omega_{kj} \underset{\rightarrow}{a.s} E_\tau(\omega_{k\tau})$ according to law of large numbers. Then we have:

$$E_\tau(\omega_{k\tau}) = E_u\left(\frac{\tau_k\Lambda_u - \tau_k u}{f(F^{-1}(\tau_k))f(F^{-1}(u))}\right),$$
$$= \int_0^{\tau_k}\frac{u - \tau_k u}{f(F^{-1}(\tau_k))f(F^{-1}(u))}du + \int_{\tau_k}^{1}\frac{\tau_k - \tau_k u}{f(F^{-1}(\tau_k))f(F^{-1}(u))}$$
$$= (1-\tau_k)\int_{-\infty}^{F-1_{(\tau k)}} F(u)du + \tau_k \int_{F-1_{(\tau_k)}}^{\infty} F(u)du = 1 + II$$

By change of variables. Define $G(s) = \int_{-\infty}^{s} F(t)dt$ and with $G(\infty) = 0$, then we have:

$$G(s) = \int_{-\infty}^{s}(s-x)f(x)dx = sF(s) - k_1(s)$$

where, $k_1(s) = \int_{-\infty}^{s} xf(x)dx. Let\, s = F^{-1}(\tau_k)$, then:

$$I + II = \frac{G(s) - S\tau_k}{f(s)} = -\frac{k_1(s)}{f(s)} = -\frac{\int_{-\infty}^{s} sf(x)dx}{f(s)} = -\frac{\int_0^{\tau_k F^{-1}(t)dt}}{f(F^{-1}(\tau_k))} = v_k$$

Especially, it is worth mentioning here that the term $\int_0^{\tau_k} F^{-1}(t)dt$ can be approximated by $F^{-1}(\tau_k/2)\tau_k$.

This completes the proof.

Proof of Theorem 2

Under H_0, we have $\sqrt{n}(\hat{\beta}_{\tau k}^{QR} - \hat{\beta}_{DQR}) \underset{\rightarrow}{D} N(0,(\omega^2(\tau_k) + 2v(\tau_k) + 1)D^{-1})$.

Consequently, for any fixed constant $t \in (0,1) T_n(t) \underset{\rightarrow}{D} X^2$ with degrees of freedom is p, the dimension of coefficient β.

In addition, for the proof of test process $\{T_n(\tau), \tau \in I\}$, we let $\beta(\tau) = \beta_\tau^{QR} - \beta_{DQR}$, $\eta^2(\tau) = \omega^2(\tau) + 2v(\tau) + 1$, then $T_n(\tau)$ can be represented as $T_n(\tau) = n\beta(\tau)D\beta(\tau)/\eta^2(\tau)$ and interpret δ as $\delta = \sqrt{n}(b - \phi)/\eta(\tau)$ for some choice of b, then according to Lemma 3.1 of GJKP and under conditions A1-A3, we have for any fixed $C > 0$:

$$\sup\{|T_n(\delta, \tau| : \|\delta\| \le C\sqrt{\log\log n}, \tau \in I\} \to 0$$

And with a variant of Theorem 1 of GJ, we can obtain the following equation:

$$\hat{\delta}_n(\tau) = \sqrt{n}(\hat{\beta}_n(\tau) - \beta(\tau))/\eta(\tau) = D_{gn}^{-1} + op(1) \qquad (7.1)$$

where, $g_n = n^{-1/2}\sum x_i\psi_\tau(u_i(\tau)), u_i(\tau) = \varepsilon_i - F^{-1}(\tau), \psi_\tau(u) = \tau - I(u<0)$. Where this representation of equation (7.1) holds uniformly on interval I. Thus $\hat{\delta}_n \underset{\rightarrow}{W} D^{-1/2}B_p$. So we have $T_n(\tau) \underset{\rightarrow}{W} Q_p^2(\tau)$ for $\tau \in I$ uniformly holds. That completes the proof of Theorem 2.

Proof of Theorem 3

For model:

$$Y = X^T\beta + (X^T\gamma)\varepsilon$$

The τ_kth conditional quantile function $Q_{\tau k}(Y|X) = X^T(\beta + \gamma c_{\tau k}) \square X^T b(\tau_k)$. Then:

$$\hat{b}(\tau_k) = \underset{\beta}{Arg\min}\sum_{i=1}^{n}\rho_{\tau k}(Y_i - X_i^T\beta)$$

According to GJ representation, for $\tau \in I \subset (0, 1)$, the following expression:

$$\sqrt{n}\left(\hat{b}(\tau) - b(\tau)\right) = \frac{1}{f(c_\tau)}G^{-1}g_n(\tau) + op(1)$$

Uniformly holds, where $g_n(\tau) = \sqrt{n}\sum x_i\psi_\tau(\varepsilon_i - c_\tau)$, $\psi_\tau(u) = \tau - I(u<0)$, $G = \frac{1}{n}X^T\Gamma X$ and $\Gamma = diag(\sigma_i)$. Thus by simple calculation, $E\left(\hat{b}\left(\tau_k\right)\right)=b\left(\tau_k\right)$ since $E(g_n(\tau)) = 0$ and $Var\left(\hat{b}\left(\tau_k\right)\right)=\frac{1}{n}\omega^2\left(\tau_k\right)G^{-1}DG^{-1}=\frac{1}{n}\omega^2\left(\tau_k\right)\Omega$ Then we have:

$$\sqrt{n}\{\{\hat{b}(\tau_k)-b(\tau_k)\}\to N(0,\omega^2(\tau_k)\Omega)$$

Thus:

$$E(\hat{\beta}_{DQR})E=\left(\frac{1}{N}k=1\sum^{N}\hat{b}(\tau_k)\right)=\frac{1}{N}\sum_{k=1}^{N}b(\tau_k)=\beta+\frac{1}{N}\sum^{N}c_{\tau k}$$

as N → ∞, we have $\frac{1}{N}\sum_{k=1}^{N}c_{\tau k}\underset{\to}{P}0,\frac{1}{N}\sum_{k=1}^{N}c_{\tau k}^2\underset{\to}{P}1$. Then

$$E(\hat{\beta}_{DQR})=\beta+o_p(1/N).$$

$$Var(\hat{\beta}_{DQR})=Var\left(\frac{1}{N}\sum_{k=1}^{N}\hat{b}(\tau_k)\right)=\frac{1}{N^2}\sum_{i=1}^{N}\sum_{i=1}^{N}Cov(\hat{b}(\tau_i),\hat{b}(\tau_j))$$
$$=\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{i=1}^{N}\omega_{ij}\right)\left(\frac{1}{n}G^{-1}DG^{-1}\right)=\frac{1}{n}\left(\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\omega_{ij}\right)\Omega$$

As with N → ∞, the term $\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\omega_{ij}\to 1$. Thus, Var $(\hat{\beta}_{DQR})=\frac{1}{n}\Omega+o_p(1/N^2)$.

Then the first equation in Theorem 3:

$$\sqrt{n}(\hat{\beta}_{DQR}-\beta)\underset{\to}{D}N(0,\Omega)$$

Holds now we proceed to prove the second expression:

$$\hat{\Sigma}_{DQR}=\frac{1}{N}\sum_{k=1}^{N}(\hat{b}(\tau_k)-\hat{\beta}_{DQR})(\hat{b}(\tau_k)-\hat{\beta}_{DQR})^T$$

Then:

$$E(\hat{\Sigma}_{DQR})E=\left(\frac{1}{N}\sum_{k=1}^{N}\hat{b}(\tau_k)-\hat{\beta}_{DQR}(\hat{b}(\tau_k)-\hat{\beta}_{DQR})^T\right)$$
$$=\frac{1}{N}\sum_{k=1}^{N}E\left(\hat{b}(\tau_k)-E(\hat{\beta}_{DQR})\right)\left(\hat{b}(\tau_k)-E(\hat{\beta}_{DQR})\right)^T+Var(\hat{\beta}_{DQR})$$
$$+2E[(\hat{b}(\tau_k)-E(\hat{\beta}))^T(E(\hat{\beta}_{DQR})-\hat{\beta}_{DQR})]$$
$$=\frac{1}{N}\sum_{k=1}^{N}\{I_k+Var(\hat{\beta}_{DQR})+II_k\}$$
$$I_k=E(\hat{b}(\tau_k)-E(\hat{\beta}_{DQR}))(\hat{b}(\tau_k)-E(\hat{\beta}_{DQR}))^TE=[(\hat{b}(\tau_k)$$
$$-\beta)(\hat{b}(\tau_k)-\beta)^T]=\gamma\gamma^Tc_{\tau k}^2G^{-1}DG^{-1}=\Sigma_Bc_{\tau k}^2+\frac{1}{n}\omega_k^2\Omega$$

$$II_k=2E[(\hat{b}(\tau_k)-E(\hat{b}))^T(E(\hat{\beta}_{DQR})-(\hat{\beta}_{DQR})]=2\{b(\tau_k)\beta^T$$
$$-E[(\hat{b}(\tau_k)\hat{\beta}_{DQR}^T]\}=2\Sigma_B+2\gamma\beta^Tc\tau_k-2E[(\hat{b}(\tau_k)\hat{\beta}_{DQR}^T]$$

Where:

$$E[\hat{b}(\tau_k)\hat{\beta}_{DQR}^T]=E\left(\hat{b}(\tau_k).\frac{1}{N}\sum_{j=1}^{N}\hat{b}(\tau_k)^T\right)$$
$$=\left(\frac{1}{N}\sum_{j=1}^{N}\omega k_j\right).\frac{1}{n}G^{-1}DG^{-1}+\frac{1}{N}\sum_{j=1}^{N}b(\tau_k)b^T(\tau_j)$$
$$=\frac{1}{N}\left(\frac{1}{N}\sum_{j=1}^{N}\omega k_j\right)\Omega+\Sigma_B+\beta\gamma^Tc\tau_k+o_p(1/N)$$

Thus $II_k=\frac{2}{n}(\frac{1}{N}\sum_j\omega k_j)\Omega+o_p(1/N)$ substitute these expressions into the above equation, we can obtain:

$$E(\hat{\Sigma}_{DQR})=\Sigma_B\left(\frac{1}{N}\sum_k c\tau_k^2\right)$$
$$+\frac{1}{n}\left(\frac{1}{N}\sum_k\omega_k^2\right)\Omega+\frac{1}{n}\Omega-\frac{2}{n}\left(\frac{1}{N}\sum_{k,j}\omega k_j\right)\Omega$$
$$=\Sigma_B=\frac{1}{n}(\theta+1)\Omega-\frac{2}{\Omega}\Omega+o_p(1/N^2)$$
$$=\Sigma_B=\frac{1}{n}(\theta-1)\Omega+o_p(1/N^2)$$

where, $\frac{1}{N}\sum_k\omega_k^2\to\theta=\int_{\in}^{1-\in}\frac{u(1-u)}{f(F^{-1}(u))}du=\int_{-a}^{a}F(t)(1-F)(t))dt$ and a is a certain constant.

To see the variance of $\hat{\Sigma}_{DQR}$ note that τ ~ U(0, 1), thus $\sqrt{n}(\hat{b}(\tau)-b(\tau))\,\square\,N(0,\frac{1}{n}\omega^2(\tau)\Omega)$. For randomly sampled {$\tau_k$, k = 1,...,N} from uniform distribution, $\{\hat{b}(\tau_k)\}_1^N$ are realizations of random variable b(τ). Then we denote:

$$m_2=E_\tau[b(\tau)-E_\tau(b(\tau))][b(\tau)-E_\tau(b(\tau))]^T=\gamma\gamma^T=\Sigma_B$$
$$m_4=E_\tau\{[b(\tau)-E_\tau(b(\tau))][b(\tau)-E_\tau(b(\tau))]^T\}^2=\Sigma_B^2E(\varepsilon^4)$$

Then $m_4-m_2^2=\lambda^2\Sigma_B^2$ where $\lambda^2=E(\varepsilon^2-1)^2$ and according to the variance of sample variance, we have:

$$V_{ar}(\hat{\Sigma}_{DQR})=Var\left[\frac{1}{N}\sum_{k=1}^{N}(\hat{b}(\tau_k)-\hat{\beta}_{DQR})(\hat{b}(\tau_k)-\hat{\beta}_{DQR})^T\right]$$
$$=\frac{1}{N}(m_4-m_2^2)+O(N^{-2})=\frac{1}{N}\lambda^2\Sigma_B^2+O(N^{-2})$$

Thus we completes the proof.

9. REFERENCES

Andrews, D.W., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica: J. Econometric Society, 61: 821-856.

Anscombe, F., 1961. Examination of residuals. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, Jun. 20-Jul. 30, University of California Press, Berkeley, Calif., pp: 1-36.

Atkinson, A.C., 1985. Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis. 1st Edn., Clarendon Press, ISBN-10: 0198533594, pp: 282.

Bickel, P., 1978. Using residuals robustly I: Tests for heteroscedasticity, nonlinearity. Ann. Stat., 6: 266-291.

Cook, R.D. and S. Weisberg, 1983. Diagnostics for heteroscedasticity in regression. Biometrika, 70: 1-10. DOI: 10.1093/biomet/70.1.1

De Long, D.M., 1981. Crossing probabilities for a square root boundary by a Bessel process. Commun. Stat. Theory Methods, 10: 2197-2213. DOI: 10.1080/03610928108828182

Fan, J., 1993. Local linear regression smoothers and their minimax efficiencies. Ann. Stat., 21: 196-216.

Fan, J. and I. Gijbels, 1996. Local Polynomial Modelling and its Applications: Monographs on Statistics and Applied Probability 66. 1st Edn., CRC Press, ISBN-10: 0412983214, pp: 360.

He, X., 1997. Quantile curves without crossing. Am. Statist., 51: 186-192. DOI: 10.1080/00031305.1997.10473959

Koenker, R., 2005. Quantile Regression. 1st End., Cambridge University Press, Cambridge, ISBN-10: 0521608279, pp: 349.

Koenker, R. and G. Bassett Jr, 1982. Robust tests for heteroscedasticity based on regression quantiles. Econometrica, 50: 43-61.

Koenker, R. and J.A.F. Machado, 1999. Goodness of fit and related inference processes for quantile regression. J. Am. Stat. Ass., 94: 1296-1310. DOI: 10.1080/01621459.1999.10473882

Koenker, R. and Q. Zhao, 1994. L-estimatton for linear heteroscedastic models. J. Nonparametric Stat., 3: 223-235. DOI: 10.1080/10485259408832584

Wicox, R.R. and H.J. Keselman, 2006. Detecting heteroscedasticity in a simple regression model via quantile regression slopes. J. Stat. Comput. Simulat., 76: 705-712. DOI: 10.1080/10629360500107923

Xiong, Tang and Tian. 2012. Estimation of heteroscedastic model via dynamica quantile regression. (Submitted).

Yu, K. and M.C. Jones, 1998. Local linear quantile regression. J. Am. Stat. Ass., 93: 228-237. DOI: 10.1080/01621459.1998.10474104

Zou, H. and M. Yuan, 2008. Composite quantile regression and the oracle model selection theory. Ann. Stat., 36: 1108-1126.