GENERALIZED QUASI-LIKELIHOOD RATIO TESTS FOR VARYING
COEFFICIENT QUANTILE REGRESSION MODELS

by

Yonggang Wang

A dissertation submitted to the faculty of
the University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2013

Approved by:

_____

Dr. Zongwu Cai

_____

Dr. Jiancheng Jiang

_____

Dr. Weihua Zhou

_____

Dr. Yongge Wang

ABSTRACT

YONGGANG WANG. Generalized quasi-likelihood ratio tests for varying
coefficient quantile regression models.
(Under the direction of DR. ZONGWU CAI)

Quantile regression models which can track the relationship of predictive vari-
ables and the response variable in specific quantiles are especially useful in applica-
tions when extreme quantiles instead of the center of the distribution are interesting.
Compared to classical conditional mean regressions, quantile regression models can
provide a more comprehensive structure of the conditional distribution of the re-
sponse variable. Also, they are more robust to skewed distributions and outliers.
Therefore, quantile regression models have been applied extensively in many ap-
plied areas. Due to its greater flexibility, a varying coefficient regression technique
has been extended to the quantile regression models recently. In this dissertation,
my aim is to propose a new test procedure, termed as generalized quasi-likelihood
(GQLR) test, to test whether all or partial coefficients are indeed constant or of
some specific functions for the varying coefficient quantile regression models. The
test statistics are constructed based on the comparison of the quasi-likelihood func-
tions under null and alternative hypotheses. The asymptotic distributions of the
proposed test statistics are also derived.

First, the functional coefficients in a varying coefficient quantile regression model
are estimated by applying local linear fitting technique with jackknife method. Then,
I construct the generalized quasi-likelihood ratio test statistics to test whether the
varying coefficients are of some specific functional forms, including two special cases:
testing whether the varying coefficients are known or unknown constants. The
asymptotic normality of the proposed test statistic is derived upon the Bahadur
representation of the estimators. I also discuss how to estimate the asymptotic
variance-covariance matrix and investigate the power of the proposed test procedures

in Chapter 2.

Secondly, I consider the similar testing procedure to test if partial coefficients in a varying coefficient quantile regression model are constant or of some specific form with other coefficients completely unspecified in Chapter 3. The corresponding generalized quasi-likelihood ratio test statistic is constructed based on comparing the quasi-likelihood functions under the null and alternative hypotheses. The asymptotic distributions of the proposed test statistics for both constancy and specific functional form are derived respectively and the power of the proposed test procedures is also investigated.

Finally, to exam the finite sample performance of all test statistics proposed in Chapters 2 and 3, Monte Carlo simulation studies are conducted respectively at the end of each chapter. I also apply the proposed test methodologies to test if the existing models in the literature used to analyze the Boston house price data are appropriate or not. The simulation results and the real example illustrate the effectiveness and practical usefulness of the proposed test statistics. Chapter 4 concludes the dissertation. I also discuss some future research topics related to this dissertation.

ACKNOWLEDGMENTS

First, I would like to gratefully and sincerely thank my Ph.D. supervisor Dr. Zongwu Cai for his excellent guidance and helps during my graduate studies at UNC Charlotte. Besides providing me with an effective guidance on my research, his enthusiasm to work also inspired me to push myself for better. This dissertation would not be possible without his guidance and encouragement. He and his family have also given me lots of help for my livings since I came to United States in 2008. I have an enormous debt to him and his family.

Appreciation also extends to Dr. Jiancheng Jiang, Dr. Weihua Zhou, and Dr. Yongge Wang for serving as valuable members in my advisory committee who give lots of precious suggestions and great support on my research. I would like to express my special thanks to Dr. Jiang for his guidance of study, suggestions in my dissertation and advices in our seminars.

I would like to thank the University of North Carolina at Charlotte, the Department of Mathematics and Statistics for their supports of my graduate study. In particular, I want to thank Dr. Yuanan Diao, Dr. Mohammad Kazemi, Dr. Shaozhong Deng, Dr. Joel Avrin for their helps. Also I would like to thank Dr. Kai Zhao of Qingdao University in China for his encouragement and suggestions.

Finally, my special gratitude goes to my wife Meijuan Jia, my parents and my brothers for their endless love, encouragement, and support throughout my life. To them I dedicate this dissertation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION AND RELATED LITERATURE

It is well known that conditional mean regression models track how the conditional expectation of the response variable depends on explanatory variables. These models are simple to fit by using least square method or maximum likelihood estimation and they are easy to interpret due to the simple model structure. However, conditional mean regression models may not be able to detect the relationship between response variable and predictive variables at a specific quantile, say the 5th or 95th quantile. For this regard, quantile regression models are in nature since they characterize how the complete distribution of the response variable depends on the predictive variables by regression at different predetermined quantiles.

Let Y be a real-valued random variable and its conditional probability distribution function of Y given X be $F(y|x) = P(Y \leq y|X = x)$. Then, for any $\tau \in (0,1)$, the τ-th conditional quantile $q_\tau(x)$ is defined as

$$q_\tau(x) = \inf\{y \in R|F(y|x) \geq \tau\}, \tag{1.1}$$

which is also called the quantile regression function. Koenker and Bassett (1978) first introduced the linear quantile regression for any $0 < \tau < 1$ and they used the so-called "check" function as the loss function to estimate the conditional quantile for any τ-th quantile:

$$q_\tau(x) = \text{argmin}_{a \in R} E\{\rho_\tau(Y - a)|X = x\}, \tag{1.2}$$

where $\rho_\tau(y) = y(\tau - I_{\{y<0\}})$ is the check function and I_A is the indicator function of

any set A. It is obvious that median regression which is well-known as least absolute deviation in the literature is a special case of quantile regression by setting $\tau = 0.5$.

Clearly, instead of focusing on conditional mean, a quantile regression model can investigate how the whole conditional distribution of Y depends on predictors X by estimating the conditional quantile function at any predetermined position (a specific quantile) of the conditional distribution of Y given X. Furthermore, it is well known that a quantile regression estimate is robust against outliers and skewed distributions. With these advantages, quantile regression models have developed swiftly during the recent years and have been widely used in many research fields, including finance, economics, medicine, biology and others. There is a vast literature about quantile regression models. Here are some examples in finance and economics. For example, Engle and Manganelli (2004) proposed the conditional autoregressive value at risk (CAVaR) by using nonlinear quantile regression estimation methods, Barnes and Hughes (2002) investigated the complete distributional impact of factors on returns of securities, Bassett and Chen (2001) applied the quantile regression method to analyze the style of a fund manager over the entire conditional distribution, and a comprehensive review of applications of quantile regression models in different areas can be found in Koenker (2005) and the references therein. Also, quantile regression can be applied to test heteroscedasticity and to construct the prediction intervals given the historical data under the stationary time series setting. For more details, the reader is referred to the papers by Koenker and Bassett (1982a, 1982b), Granger, White and Kamstra (1989), Efron (1991), Koenker and Zhao (1996), Taylor and Bunn(1999), Koenker and Xiao (2002) and the references therein.

By modeling any predetermined quantiles, quantile regressions can be used to measure the effect of covariates not only in the center of the conditional distribution but also in some extreme quantiles, such as the 5th and 95th quantiles. Without as-

suming that quantiles are related to covariate X in a specific form, one can estimate the interesting quantile directly by using model (1.2). In contrast to the conditional mean model, the effects of a covariate in a conditional quantile model can be quite different for the upper and lower tails. One famous and interesting example in the literature is that the Democrats claimed that "the rich got richer and the poor got poorer" during the Republican administrations in the 1992 presidential selection in the United States. How to verify whether this claim is valid or do people have enough evidence to reject this claim? Conditional mean regression models can not answer this question directly but quantile regression can. One can compute 90th or higher quantile functions of the number of people in the high-salary category and 10th or lower quantile functions of people in the low-salary category respectively. An increasing 90th or higher quantile functions and a decreasing 10th or lower quantile functions will support the claim. The readers are referred to Figure 6.4 in Fan and Gijbels (1996) for detail. Indeed, Rose (1992) analyzed the data for the 1979-1989 period and showed that the 10th quantile and the 90th quantile of the family income indeed display two opposite trends over time. Similarly, in survival analysis, it is of great interest to study the effect of a covariate on high risk individuals as well as the effect on different risk levels (different quantiles). Another good example can be found in Hendricks and Koenker (1992), in a study of consumer demand for electricity, heavy users responded much more drastically to weather and time variation than average users.

One important application of quantile regression in finance is the implementation of the so-called value-at-risk (VaR). As one of the important modern risk measuring techniques, VaR is indeed the lowest quantile of the potential losses that can occur with a given portfolio during a specified time period and it measures the worst expected loss under normal market conditions over a specific time interval for a given confidence level. There is a huge literature in the area of VaR. The details

can be found in J.P. Morgan (1995), Duffile and Pan (1997), Engle and Manganelli (1999), Tsay (2000, 2002) and Khindanova and Rachev (2000) and so on.

1.1 Linear Quantile Regression

For the given sample $\{(X_i, Y_i), i = 1, ..., n\}$, Koenker and Bassett (1978) suggested that the linear conditional quantile function $q_\tau(x) = Q_y(\tau|x) = x'\beta(\tau)$ can be estimated through

$$\hat{\beta}(\tau) = \mathrm{argmin}_{\beta \in R} \sum_{i=1}^{n} \rho_\tau(y_i - X_i'\beta(\tau)) \tag{1.3}$$

for any $\tau \in (0, 1)$ and $\rho_\tau(y) = y(\tau - I_{y<0})$ is the check function, which is a special case of the loss function discussed in Huber (1964), which considered the minimization problem to obtain the M-estimator for β by choosing some suitable loss function. If setting $\tau = \frac{1}{2}$, the right hand side of equation (1.3) becomes the sum of absolute deviation errors and it is called median regression or LAD regression.

There is a huge literature devoted to this model for both cross-sectional and time series data; see Duffie and Pan (1997), Koenker (1978, 1982a, 2000), Tsay (2000), Koenker and Hallock (2001) and the references therein. Suppose that the τ-th conditional quantile function of Y given X can be written as

$$Q_y(\tau|x) = x^T\beta(\tau). \tag{1.4}$$

The following conditions are needed to ensure that $\hat{\beta}(\tau)$ in (1.3) is consistent.

Condition A1: There exists $d > 0$ such that

$$\liminf_{n \to \infty} \inf_{||u||=1} n^{-1} \sum I(|x_i^T u| < d) = 0$$

Condition A2: There exists $D > 0$ such that

$$\liminf_{n \to \infty} \inf_{||u||=1} n^{-1} \sum (x_i^T u)^2 \le D$$

Assuming that the design matrix X satisfies the conditions in A1 and A2, El Bantli and Hallin (1999) presented the necessary and sufficient condition for consistency of estimator (i.e. $\hat{\beta}(\tau) \to \beta(\tau)$) and the conditional distribution functions F_{ni} of $Y_i, i = 1, 2, ...,$ satisfy

Condition B. $\sqrt{n}(a_n(\varepsilon) - \tau) \to \infty$ and $\sqrt{n}(\tau - b_n(\varepsilon)) \to \infty$ with $a_n(\varepsilon) = n^{-1} \sum F_{ni}(x_i^T \beta(\tau) - \varepsilon)$ and $b_n(\varepsilon) = n^{-1} \sum F_{ni}(x_i^T \beta(\tau) + \varepsilon)$.

Having established the result of consistency of estimators, one may be interested in the asymptotic distribution for the quantile regression estimators.

Condition C1. The distribution of X, $F(x)$, has continuous and strictly positive density, $f(x)$, for all x such that $0 < F(x) < 1$.

Condition C2. $\lim_{n \to \infty} n^{-1} \sum_{t=1}^{n} X_t X_t^T = D$ is a positive definite matrix.

Under the above conditions, Koenker and Bassett (1982b) established the following theorem.

Theorem 1.1: Let $(\hat{\beta}(\tau_1)), ..., \hat{\beta}(\tau_m)))$ with $0 < \tau_1 < ... < \tau_m < 1$ denote a sequence of regression quantiles. Under Conditions C1 and C2, as $n \to \infty$,

$$\sqrt{n}(\hat{\beta}(\tau_1) - \beta(\tau_1), ..., \hat{\beta}(\tau_m) - \beta(\tau_m))$$

converges in law to an mp-variate Gaussian random vector with zero mean and covariance matrix $\Omega(\tau_1, ..., \tau_m, F) \otimes D^{-1}$, where $\otimes$ denotes the Kronecker product

and Ω has the typical (i, j)th element as

$$\omega_{ij} = \frac{\min(\tau_i, \tau_j) - \tau_i \tau_j}{f(F^{-1}(\tau_i)) f(F^{-1}(\tau_i))}.$$

Powell (1986) introduced the quantile regression to the Tobit model, which can be written in the following form

$$Y_t = \max(0, x_t^T \beta_0 + \varepsilon_t), \quad t = 1, ..., n,$$

and Koenker and Zhao (1996) used the quantile regression method to an autoregressive conditional heteroscedastic (ARCH) type model as

$$Y_t = (\gamma_0 + \gamma_1 |Y_{t-1}| + ... + \gamma_q |Y_{t-q}|) \epsilon_t,$$

where $0 < \gamma_0 < \infty, \gamma_1, ..., \gamma_q \geq 0$, and $\{\epsilon_t\}$ are i.i.d. random variables with distribution $F_\epsilon(\cdot)$ and density function $f_\epsilon(\cdot)$.

1.2 Nonlinear Quantile Regression

Linear quantile regression models are easy to fit and interpret, but they may not be flexible enough to capture the underlying complex dependence structure of the quantile of the response variable and its covariates. At this stage, it is of considerable interest to investigate conditional quantile models which are nonlinear in parameters: $Q_y(\tau|x) = g(x, \beta(\tau))$ with known form of $g(\cdot, \cdot)$. Nonlinear quantile regression models are more general, including the linear quantile regression models as a special cases. The nonlinear quantile regression estimator is defined as following:

$$\hat{\beta}_n(\tau) = \operatorname{argmin}_{b \in \mathcal{B}} \sum \rho_\tau(y_i - g(x_i, b)),$$

where $\mathcal{B} \subset R^p$ is compact set. As discussed in Koenker (2005), we need some further conditions besides Condition C1 to obtain the asymptotic distribution of the estimator.

Condition D1. There exist constants k_0, k_1, and n_0 such that, for $\beta_1, \beta_2 \in \mathcal{B}$ and $n > n_0$,

$$k_0||\beta_1 - \beta_2|| \le (n^{-1} \sum_{i=1}^{n} (g(x_i, \beta_1) - g(x_i, \beta_2))^2)^{1/2} \le k_1||\beta_1 - \beta_2||$$

Condition D2. There exist positive definite matrixes D_0 and $D_1(\tau)$ such that, with $\dot{g}_i = \partial g(x_i, \beta)/\partial\beta|_{\beta=\beta_0}$

$$i) \quad \lim_{n \to \infty} n^{-1} \sum \dot{(g)}_i \dot{g}_i^T = D_0$$

$$ii) \quad \lim_{n \to \infty} n^{-1} \sum f_i(\xi_i)\dot{(g)}_i \dot{g}_i^T = D_1(\tau)$$

$$iii) \quad \max_{i=1,...,n} ||\dot{g}_i||/\sqrt{n} \to 0$$

Under these conditions, one can derive the Bahadur representation of quantile regression estimator as

$$\sqrt{n}(\hat{\beta}_n(\tau) - \beta_0(\tau)) = D_1^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{g}_i \psi_\tau(u_i(\tau)) + o_p(1),$$

where $u_i(\tau) = y_i - g(x_i, \beta_0(\tau))$. The Bahadur representation has been investigated in an extensive way in the literature for the quantile regression estimator since it can represent the complicated nonlinear estimator as a simple normalized sum of i.i.d variables. Upon the Bahadur representation, the asymptotical normal distribution of proposed estimators can be derived easily. Consequently, we have

$$\sqrt{n}(\hat{\beta}_n(\tau) - \beta_0(\tau)) \to N(0, \tau(1 - \tau)D_1^{-1}D_0 D_1^{-1}).$$

See Koenker (2005) for details.

Due to fast developing computing technology in computer science, the computing speed of complex model structure becomes available to researchers. Nonparametric and semiparametric quantile regression models which need great computing power have attracted a great deal of research attentions due to their greater flexibility. The literature on nonparametric and semiparametric quantile models is large. The readers are referred to the papers by, to name just a few, Chaudhuri (1991), He and Shi (1996), He and Ng (1999), Cai (2002), Yu and Lu (2004), Cai and Xu (2008), Cai, Gu and Li (2009) and the references therein.

As we all known that, in additional to difficult practical implementation, nonparametric quantile regression with multivariate covariates also suffers from the "curse of dimensionality". Many dimension-reduction techniques have been adopted for quantile regression to deal with this problem, such as additive model, single index model and varying coefficient quantile regression models. Honda (2004) and Cai and Xu (2008) considered the varying coefficient model for time series data, Wu, Yu and Yu (2010) investigated the single index model for quantile regression, and He, Ng and Portnoy (1998), He and Ng (1999) and Horowitz and Lee (2005) discussed the additive quantile regression model for the i.i.d data.

1.3 Varying Coefficient Quantile Regression Models

A varying coefficient regression model is a useful and natural extension of a classical linear regression model. The modeling technique for such a model has been extensively discussed in the seminal work of Cleveland, Grosse and Shyu (1991) and Hastie and Tibshirani (1993). The varying coefficient models assume the following conditional mean structure:

$$Y = \sum_{j=1}^{p} a_{j,m}(U)X_j + \varepsilon, \tag{1.5}$$

where $a_{j,m}(\cdot)$ denotes the varying coefficient function for the mean regression. As argued by Hastie and Tibshirani (1993), this model can avoid the "curse of dimensionality" and reduce the modeling bias significantly and it is also easily interpreted by allowing the coefficients to depend on some smooth variables.

Indeed, Cai (2010) argued that a functional-coefficient model given in (1.5) has several strengths, at least including following three advantages. First, it can be actually a good approximation to a general fully nonparametric model. Secondly, it has an ability to capture heteroscedasticity. Finally, it can be used as a tool to study covariate adjusted regression for situations where both predictors and response in a regression model are not directly observable, but are contaminated with a multiplicative factor that is determined by the value of an unknown function of an observable covariate (confounding variable).

One simple approach to estimate the coefficients is to use the local linear or local polynomial estimation technique. To maintain the simplicity of the dissertation, I will focus on the local linear technique in what follows as recommended by Fan and Gijbels (1996). All are similar for local polynomial estimation methods except complex notations. For each given point u_0, by Taylor's expansion we can approximate the function locally

$$a_{j,m}(U_i) \approx a_{j,m} + b_{j,m}(U_i - u_0)$$

for any U_i in a neighborhood of u_0, a given grid point within the domain of U_i, where $b_{j,m}$ is the first derivative of $a_{j,m}(u)$. Then, we can minimize the following locally weighted least squares

$$\sum_{i=1}^{n}(Y_i - \sum_{j=1}^{p} x_{ij}(a_{j,m} + b_{j,m}(U_i - u_0)))^2 K_h(U_i - u_0)$$

to get the estimate of the functional coefficient $a_{j,m}(\cdot)$. By moving u_0 along with the domain of U_i, the estimate $\hat{a}_j(u_0)$ of the entire curve $a_{j,m}(u_0)$ is be obtained.

Cai and Xu (2008) adopted the varying coefficient modeling approach for quantile regression

$$q_\tau(U, X) = \sum_{k=0}^{p} a_{k,\tau}(U) X_k,$$

where $q_\tau(U, X)$ is the τth conditional quantile of Y given U and X, to analyze dynamic time series data with α-mixing conditions. Their model covers many familiar quantile regression models, including the quantile autoregressive model proposed by Koenker and Xiao (2004) and heteroscedastic linear models considered by Koenker and Zhao (1996). Cai and Xu (2008) applied the above local linear technique to estimate the coefficients of quantile regression model (see (2.6) later for details) and derived the Bahadur representation for the estimators.

Denote $\Omega(u_0) = E[X_t X_t' | U_t = u_0]$ and $\Omega^*(u_0) = E[X_t X_t' f_{y|u,x}(q_\tau(u_0, X_t)) | U_t = u_0]$, where $f_{y|u,x}(y)$ is the conditional density of Y given U and X. Let $f_u(u)$ present the marginal density of U. Cai and Xu (2008) imposed the following assumptions for the asymptotic distribution of the estimators.

Assumption A:

(A1) $A(u)_\tau = (a_{0,\tau}(u), \cdots, a_{p,\tau}(u))'$ is continuously twice differentiable in a neighborhood of u_0 for any given grid point u_0.

(A2) $f_u(u)$ is continuous, and $f_u(u_0) > 0$.

(A3) $f_{y|u,x}(y)$ is bounded and satisfies the Lipschitz condition.

(A4) The Kernel function $K(\cdot)$ is symmetric and has a compact support, say$[-1,1]$.

(A5) (U_t, X_t, Y_t) is a strictly α-mixing stationary process with mixing coefficient $\alpha(t)$ satisfying $\sum_{t \geq 1}^{\infty} t^l \times \alpha^{(\delta-2)/\delta}(t) < \infty$ for some positive real number $\delta \geq 2$ and $l > (\delta - 2)/\delta$.

(A6) $E||X_t||^{2\delta^*} < \infty$ with $\delta^* > \delta$.

(A7) $\Omega(u_0)$ is positive definite and continuous in a neighborhood of u_0.

(A8) $\Omega^*(u_0)$ is positive definite and continuous in a neighborhood of u_0.

(A9) The bandwidth h satisfy $h \to 0$ and $nh \to \infty$.

(A10) $f(u, v|x_0, x_s; s) \le M < \infty$ for $s \ge 1$, where $f(u, v|x_0, x_s; s)$ is the conditional density of (U_0, U_s) given $(X_0 = x_0, X_s = x_s)$.

(A11) $n^{1/2 - \delta/4} h^{\delta/\delta^* - 1/2 - \delta/4} = O(1)$.

Under these assumptions, they derived the Bahadur representation (see (2.7) later for details) as follows

$$\sqrt{nh}(\hat{A}_\tau(U_t) - A_\tau(U_t)) = \frac{(\Omega^*(U_t))^{-1}}{\sqrt{nh}f_u(U_t)} \sum_{i \neq t} \varphi_\tau(Y_i^*)X_i K(\frac{U_i - U_t}{h}) + o_p(1), \quad (1.6)$$

where $\hat{A}_\tau(\cdot)$ is the local linear estimate of $A_\tau(U_t)$ (see (2.6) later for details) and $\varphi_\tau(u) = \tau - I_{\{u < 0\}}$. Then, they derived the following theorem about the asymptotical normal distribution of estimators.

Theorem 1.2: Under Assumption A, we have the following asymptotic normality,

$$\sqrt{nh}\left[\hat{A}_\tau(u_0) - A_\tau(u_0) - \frac{h^2}{2}A_\tau''(u_0)\mu_2 + o_p(h^2)\right] \to N(0, \tau(1 - \tau)\nu_0\Sigma(u_0)), \quad (1.7)$$

where $\Sigma(u_0) = [\Omega^*(u_0)]^{-1}\Omega(u_0)[\Omega^*(u_0)]^{-1}/f_u(u_0)$, $\mu_j = \int u^j K(u)du$ and $v_j = \int u^j K^2(u)du$.

For this varying coefficient quantile regression model, some great interesting inference questions arise naturally in practice. One may be interested in testing whether all or partial coefficients are really varying, or of some specific functional form, and if certain components of covariates X are statistically insignificant. This

leads to the following test hypotheses:

$$H_0 : A_\tau(u) = A_{0,\tau}(u) \leftrightarrow H_a : A_\tau(u) \neq A_{0,\tau}(u), \qquad (1.8)$$

or

$$H_0 : A_{1,\tau}(u) = A_{10,\tau}(u) \leftrightarrow H_a : A_{1,\tau}(u) \neq A_{10,\tau}(u), \qquad (1.9)$$

where $A_{1,\tau}(u)$ are some coefficients in $A_\tau(u)$ and $A_{0,\tau}(u)$ and $A_{10,\tau}(u)$ may be either a constant vector or a functional vector. For varying coefficient mean regression models (by changing τ to be m in the above hypotheses given in (1.8) and (1.9)), Cai, Fan and Yao (2000) and Fan, Zhang and Zhang (2001) proposed the so-called generalized likelihood ratio (GLR) test which was based on the likelihood ratio under the null hypothesis and the alternative hypothesis. Cai, Fan and Yao (2000) tested for such hypothesis by using bootstrap based method, while Fan, Zhang and Zhang (2001) considered the test based on the asymptotic result. But for the varying coefficient quantile regression models, to the best of my knowledge, there is no paper in the literature to develop the similar test procedure. In this dissertation, I develop some new test procedure, termed as generalized quasi-likelihood ration (QGLR) test, to check whether coefficients or partial coefficients are of constant or specific functional form. Before I describe the proposed test procedure for these hypothesis given in (1.8) and (1.9), I will give a brief review about inference problem for quantile regression models.

1.4 Inference for Quantile Regression Model

The classical theory of linear regression model assumes that the slope coefficients of distinct quantile regressions are identical. This assumption implies that conditional quantile functions are all parallel to each other for different quantiles. However, different slope estimates often vary greatly across quantiles in real applica-

tions as discussed above. So, to test the hypothesis of equality of slope parameters across quantiles is one of fundamental inference problems in quantile regression. Consider the two-sample model

$$Y_i = \alpha_1 + \alpha_2 x_i + u_i,$$

where $x_i = 0$ for n_1 observations in the first sample and $x_i = 1$ for n_2 observations in the seconde sample. Koenker and Bassett (1982a) proposed the Wald-type test statistic

$$T_n = (\hat{\alpha}_2(\tau_2) - \hat{\alpha}_2(\tau_1))/\hat{\sigma}(\tau_1, \tau_2)$$

to test the equality of the slop parameters across quantiles τ_1 and τ_2 with

$$\sigma^2(\tau_1, \tau_2) = \left[\frac{\tau_1(1-\tau_1)}{f^2(\xi_1)} - 2\frac{\tau_1(1-\tau_2)}{f(\xi_1)f(\xi_2)} + \frac{\tau_2(1-\tau_2)}{f^2(\xi_2)}\right]\left[\frac{n}{nn_2 - n_2^2}\right]$$

and $\xi_i = F^{-1}(\tau_1)$. For the more general linear hypothesis

$$H_0 : R\zeta = r,$$

where $\zeta = (\beta(\tau_1)^T, ..., \beta(\tau_m)^T)^T$. Koenker and Bassett (1982b) proposed the test statistic

$$T_n = n(R\hat{\zeta} - r)^T[RV^{-1}R^T]^{-1}(R\hat{\zeta} - r),$$

where V_n is the mp×mp matrix with ijth block

$$V_n(\tau_i, \tau_j) = [\tau_i \wedge \tau_j - \tau_i\tau_j]H_n(\tau_i)^{-1}J_n(\tau_i, \tau_j)H_n(\tau_j)^{-1}.$$

The test statistic is asymptotically χ_q^2 under H_0, where q is the rank of the matrix R.

The Wald test can check whether it is plausible that certain linear restrictions hold. However, the score approach can detect a direction to move in the space of alternative hypotheses that leads to more plausible estimate. Consider the model

$$Q_{y_i}(\tau|x_i, z_i) = x_i^T \beta(\tau) + z_i^T \varsigma(\tau)$$

with the associated null hypothesis,

$$H_0 : \varsigma(\tau) = 0 \leftrightarrow H_a : \varsigma_n(\tau) = \varsigma_0(\tau)/\sqrt{n}$$

Then the score test statistic can be defined as

$$T_n = S_n^T M_n^{-1} S_n / A^2(\varphi),$$

where $S_n = n^{-1/2}(Z - \hat{Z})^T \hat{b}_n)$, $\hat{Z} = X(X^T \Psi X)^{-1} X^T \Psi Z$, $\Psi = diag(f_i(Q_{Y_i}(\tau|x_i, z_i)))$, $M_n = (Z - \hat{Z})(Z - \hat{Z})^T/n$, and $\hat{b}_n = \int_0^1 \hat{a}_n(s)d\varphi(s)$. Koenker and Machado (1999) showed that T_n has a limiting central χ_q^2 distribution under the null hypothesis with some mild conditions on the score function φ, where q denotes the dimension of the parameter ζ.

It is also natural to investigate analogs of the likelihood ratio test as well as the Wald and score test for quantile regression models. Koenker and Bassett (1982b) had proposed the test statistic

$$L_n = 8(\tilde{V}(\frac{1}{2}) - \hat{V}(\frac{1}{2}))/s(\frac{1}{2})$$

to test the hypothesis $H_0 : R\beta = r$ for the median regression in the i.i.d-error linear model, with $\hat{V}(\tau) = \min_{\{b \in R^p\}} \sum \rho_\tau(y_i - x_i'b)$ and $\tilde{V}(\tau) = \min_{\{b \in R^p | R\beta = r\}} \sum \rho_\tau(y_i - x_i'b)$ under null or alternative hypothesis respectively. It was shown that L_n is

asymptotically χ_q^2, where $q = \text{rank}(R)$. One can easily realize that this statistic is related to a likelihood ratio. The same approach can be extended to other quantile regression models. Define $\hat{V}(\tau)$ and $\widetilde{V}(\tau)$ as above, and let $\hat{\sigma}(\tau) = \hat{V}(\tau)/n$ and $\widetilde{\sigma}(\tau) = \widetilde{V}(\tau)/n$. Then, the for the τth quantile the test statistic is defined as

$$L_n(\tau) = \frac{2}{\lambda^2(\tau)s(\tau)}[\widetilde{\sigma}(\tau) - \hat{\sigma}(\tau)].$$

As in Koenker (2005), we can call it as quasi-likelihood-ratio test or ρ-test by following similar terminology in Ronchetti (1985).

Motivated by these testing methodologies, I propose the quasi-likelihood ratio test statistic which is constructed based on the comparison of the quasi-likelihood functions under null and alternative hypotheses to test for varying coefficient quantile regression models about the form of coefficients or partial coefficients. In this dissertation, I focus on two testing hypotheses

$$H_0 : A_\tau(u) = A_{0,\tau}(u) \leftrightarrow H_a : A_\tau(u) \neq A_{0,\tau}(u),$$

where $A_{0,\tau}(u)$ is a vector of constants or some specific functions and

$$H_0 : A_{1,\tau}(u) = A_{10,\tau}(u) \leftrightarrow H_a : A_{1,\tau}(u) \neq A_{10,\tau}(u),$$

where $A_{1,\tau}(u)$ is a vector of partial coefficients in the varying coefficient quantile regression model with other coefficients completely unspecified and $A_{10,\tau}$ is a vector of constants or some specific functions.

1.5 Overview

The rest of this dissertation is organized as follows. In Chapter 2, I discuss the estimation of coefficients of varying coefficient quantile regression models by using

jackknife method and local fitting technique and then derive the Bahadur representation of the proposed estimators. Furthermore, I propose some new test statistics, termed as generalized quasi-likelihood ratio test, to testing if varying coefficients for varying coefficient quantile regression model are constant or of some specific functional form. The test statistics are constructed based on the comparison of the quasi-likelihood under null and alternative hypotheses respectively. The asymptotic distribution of the test statistics is derived. In order to evaluate the finite sample performance of the proposed methods, Monte Carlo studies are conducted and a real application of test procedure to Boston house price data is then reported to highlight the proposed test procedures.

In Chapter 3, I consider a more general case. For varying coefficient quantile regression models, one may be interested in testing whether some of varying coefficients are of some functional form or constants with other coefficients completely unknown. I construct similar generalized quasi-likelihood ratio test statistics to test such hypotheses. I construct the quasi-likelihood function under the null hypothesis by using the semiparametric estimators which are proposed in Cai and Xiao (2012). Again, I derive the asymptotic theory for such test statistics. Also, I conduct Monte Carlo simulation examples to show the finite sample performance of the proposed methods, and finally an application to Boston house price data is discussed to show the effectiveness of the testing methodology.

Chapter 4 concludes the dissertation and discuss some possible future research directions. The detailed proofs of the main results in each chapter are relegated to the last section of the corresponding chapter.

CHAPTER 2: GENERALIZED QUASI-LIKELIHOOD RATIO TEST OF THE
COEFFICIENTS FOR VARYING COEFFICIENTS QUANTILE REGRESSION
MODEL

This chapter mainly discusses the testing hypotheses about whether the coefficients of varying coefficient quantile regression models are of some specific functional form or constants. The simulation studies and a real application of the proposed test procedure are also reported at the end of this chapter.

2.1   Introduction and Motivation

For varying coefficient quantile regression models, after fitting the varying coefficient quantile regression models by using either local linear technique, sieve or penalized likelihood methods, one great interesting inference problem arises naturally is to check whether the varying coefficients are of some specific functional form. This is equivalent to the following hypotheses:

$$H_0 : A_\tau(u) = A_{0,\tau}(u) \quad \text{versus} \quad H_a : A_\tau(u) \neq A_{0,\tau}(u), \tag{2.1}$$

where $A_{0,\tau}(u)$ is a vector of known functionals. One special case of (2.1) is that $A_0(u)$ is a vector of constants. Then, the test hypothesis becomes to checking whether the varying coefficients are indeed varying. That is equivalent to

$$H_0 : A_\tau(u) = A_{0,\tau} \quad \text{versus} \quad H_a : A_\tau(u) \neq A_{0,\tau}, \tag{2.2}$$

where $A_{0,\tau}$ is a vector of known or unknown constants. When $A_{0,\tau}$ is a vector of unknown constants, the hypothesis (2.2) may be more interesting since people may

care about whether the varying coefficients are indeed constant without knowing specific values. These hypothesis testing is equivalent to the model assessment against the liner quantile regression model. The model assessment is very important since a linear quantile model, which is simple to implement and easy to interpret, is more preferred unless a varying coefficient model is necessary for the data or underlying structure. Recently, Cai, Fan and Yao (2000) discussed how to construct the test statistic for these hypotheses based on the generalized likelihood ratio test for a varying coefficient mean regression model as in (1.5). He and Zhu (2003) and Horowitz and Spokoiny (2002) considered general lack-of-fit tests for linear quantile regression model. Both tests are consistent for any fixed alternative.

The likelihood ratio type test was proposed by Cai, Fan and Yao (2000) and studied extensively by Fan, Zhang and Zhang (2001) for the hypothesis testing problems formulated in (2.1) and (2.2) for the conditional mean regression models in (1.5). Recall that the likelihood ratio statistic is constructed for the testing problems in (2.1) and (2.2) for conditional mean regression model in (1.5), described briefly as follows. Denote $\hat{A}_m(U)$ is the corresponding nonparametric estimator $A_m(U)$. Then, the statistic is defined as follows.

$$\lambda_n = \mathscr{L}_n(H_a) - \mathscr{L}_n(H_0) = \frac{n}{2} \log \frac{RSS_0}{RSS_a} \approx \frac{n}{2} \frac{RSS_0 - RSS_a}{RSS_a}, \qquad (2.3)$$

where $\mathscr{L}_n(H_a)$ is the log-likelihood under H_a with unknown regression function replaced by a reasonable nonparametric regression estimator, $RSS_a = \sum_{k=1}^{n}(Y_k - \hat{A}_m^T(U_k)X_k)^2$ and $RSS_0 = \sum_{k=1}^{n}(Y_k - \hat{A}_{0,m}^T(U_k)X_k)^2$, where $\hat{A}_{0,m}(U_k)$ is the true or estimated value of coefficients under H_0. Fan, Zhang and Zhang (2001) studied the asymptotic properties of the proposed test statistic in (2.3) under several situations.

Motivated by Cai, Fan and Yao (2000) and Fan, Zhang and Zhang (2001), for the varying coefficients quantile regression models, by taking the loss function as

the check function instead of the sum of squared errors, I propose the similar test statistic for the testing problems in (2.1) and (2.2) as follows

$$
\begin{aligned}
T_n &= \mathscr{L}(H_a) - \mathscr{L}(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{j=0}^{p} \hat{a}_{k,\tau}(U_t)X_{k,t} \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{k=0}^{p} \hat{a}_{k0,\tau}(U_t)X_{k,t} \right), (2.4)
\end{aligned}
$$

where $\mathscr{L}(H_a) = \sum_{t=1}^{n} \rho_\tau(Y_t - \sum_{k=0}^{p} \hat{a}_{k,\tau}(U_t)X_{k,t})$, and $\hat{a}_{k,\tau}(\cdot)$ is the nonparametric estimate of $a_{k,\tau}(\cdot)$ under the alternative hypothesis, and $\hat{a}_{k0,\tau}(\cdot)$ is the true or estimated value of $a_{k0,\tau}(\cdot)$ under H_0.

Remark 2.1: Linear quantile regression estimation is adopted when $a_{k0,\tau}$ is unknown constant under H_0.

The quasi-likelihood ratio test considered by Koenker (2005) for linear quantile models can be re-expressed as the following test statistic

$$
L_n = 8(\tilde{V}(\frac{1}{2}) - \hat{V}(\frac{1}{2}))/s(\frac{1}{2})
$$

where $\hat{V}(\tau) = \min_{\{b_\tau \in R^p\}} \sum \rho_\tau(y_i - x_i'b_\tau)$ and $\tilde{V}(\tau) = \min_{\{b_\tau \in R^p | Rb_\tau = r\}} \sum \rho_\tau(y_i - x_i'b_\tau)$. As elaborated in Komunjer (2005), $\ell(H) = \sum_{t=1}^{n} \rho_\tau \left(y_t - \sum_{k=0}^{p} a_{k,\tau}(U_t)X_{k,t} \right)$ can be regarded as the negative logarithm of quasi-likelihood, described below.

Definition 2.1: (Komunjer (2005)) A family of probability measure on $\mathscr{R}$ admitting a density φ_t^α indexed by a parameter η, $\eta_t \in M_t$, $M_t \subset \mathscr{R}$, is called tick-exponential of order α, $\alpha \in (0,1)$, if and only if:

(i) for $y \in \mathscr{R}$,

$$
\varphi_t^\alpha(y,\eta) = \exp\left(-(1-\alpha)\right)[a_t(\eta) - b_t(y)]1(y \le \eta)) + \alpha[a_t(\eta) - c_t(y)]1(y > \eta),
$$

where $a_t : M_t \to \mathscr{R}$ is continuous differentiable and $b_t : \mathscr{R} \to \mathscr{R}$ and $c_t : \mathscr{R} \to \mathscr{R}$ are $\mathscr{F}_t$-measurable; the function a_t, b_t and c_t are such that for $\eta \in M_t$;

(ii) φ_t^α is a probability density, i.e. $\int_R \varphi_t^\alpha(y, \eta)dy = 1$;

(iii) η is the α-quantile of φ_t^α, i.e. $\int_{-\infty}^\eta \varphi_t^\alpha dy = \alpha$.

From the definition, φ_t^α assigns different slopes proportional to $1 - \alpha$ and α, linear-exponential "by parts" for a given value of probability α. Gourieroux, Monfort and Renault (1987) obtained an alternative expression for φ_t^α, given by

$$\varphi_t^\alpha(y, \eta) = \exp\{g_t(y) - (1 - \alpha)[a_t(\eta) - d_t(y)]1(y \leqslant \eta) + \alpha[a_t(\eta) - d_t(y)]1(y > \eta)\}$$

by setting $d_t(y) = (1 - 2\alpha)^{-1}[(1 - \alpha)b_t(y) - \alpha c_t(y)]$ and $g_t(y) = \alpha(1 - \alpha)(2\alpha - 1)^{-1})[b_t(y) - c_t(y)]$. As discussed in Komunjer (2005), by using the special function for $a(t), b(t)$ and $c(t)$, i.e. $a_t(\eta) = [\frac{1}{\alpha(1-\alpha)}]\eta$ and $b_t(y) = c_t(y) = [\frac{1}{\alpha(1-\alpha)}]y$, the function $ln\varphi_t^\alpha$ is proportional to $t_\alpha(y, \eta) = (\alpha - 1(y \leq \eta))(y - \eta)$ which is the "check" function. Therefore,

$$\ell(H) = \sum_{t=1}^{n} \rho_\tau \left(y_t - \sum_{k=0}^{p} a_{k,\tau}(U_t)X_{k,t} \right)$$

can be regarded as the negative log of quasi-likelihood.

According to the above discussion, test statistic in (2.4) is indeed the generalized quasi-likelihood ratio (GQLR) test statistic. I will use the term generalized quasi-likelihood ratio test in what follows in the rest of this dissertation.

The rest of this chapter is organized as follows. At first I estimate the coefficients of the quantile regression model by using local fitting technique. To perform the testing hypotheses (2.1)and (2.2), I present the generalized likelihood ratio test statistic and derive the asymptotic properties of such test statistics in the following subsections. Furthermore, I investigate the power of the proposed test procedure. I also conduct Monte Carlo simulation studies and analyze Boston House Price data

to illustrate the effectiveness of the proposed methodologies.

2.2   Estimation of the Regression Coefficients

The varying coefficient quantile regression model takes the form

$$q_\tau(U_t, X_t) = \sum_{k=0}^{p} a_{k,\tau} X_{tk} = A_\tau(U_t)^T X_t = A(U_t)^T X_t, \qquad (2.5)$$

where $U_t \in R^d$ is called the smoothing variable and $X_t = (X_{t0}, X_{t1}, ..., X_{tp})'$ with $X_{t0} = 1$ are i.i.d observations, $A(U_t) = A_\tau(U_t) = (a_{0,\tau}(U_t), a_{2,\tau}(U_t), ..., a_{p,\tau}(U_t))^T$ are smooth coefficient functions which might be some function of $X_{t0},...,X_{tp}$ or time or some other exogenous variables. Without loss of generality, I consider only the case in which U_t in (2.5) is one dimensional ($d = 1$). For multivariate U_t, the modeling procedure and the related theory for the univariate case continue to hold but more complicated notations are involved and for simplicity, we drop τ from $\{a_{k,\tau}(\cdot)\}$ in what follows.

To estimate the coefficient functions $A(\cdot)$, I apply the local fitting technique with the jackknife (leave one out) method as follows. Assume $A(U)$ has a continuous first derivative. For U_i in a neighborhood of U_t, one can apply Taylor expansion to approximate $A(U_i)$ as

$$A(U_i) \approx \beta_0 + (U_i - U_t)\beta_1,$$

where $\beta_0 = A(U_t)$ and $\beta_1 = A'(U_t)$ is the first derivative of $A(U_t)$. The jackknife method is to use all observations excepts the tth observation in estimating $A(U_t)$. Then, the locally weighted loss function is given by

$$\sum_{i\neq t}^{n} \rho_\tau(Y_i - X_i(\beta_0 + (U_i - U_t)\beta_1))K_h(U_i - U_t), \qquad (2.6)$$

where $K(\cdot)$ is a kernel function, $K_h(x) = 1/hK(1/h)$, and $h = h_n$ is a sequence of

positive numbers tending to zero, which controls the amount of smoothing used in estimations. We can get the local linear estimate of $A(U_t)$, denoted by $\hat{A}_{-t} = \hat{\beta}_0$, by minimizing the above locally weighted loss function with respect to β_0 and β_1. Note that for any given U_t, the estimate $\hat{A}_{-t}(U)$ is independent of the tth observation $\{Y_t, X_t\}$, and $\hat{A}'(U_t) = \hat{\beta}_1$, the local linear estimate of the first derivative $A'(U_t)$. If taking U_i as a grid point u_0, we obtain the local linear estimate of $A(u_0)$. By moving u_0 along with the real line, one can estimate the entire curve $A(\cdot)$.

Local linear fitting can be easily implemented by modifying the existing programs for a linear quantile model slightly. For example, the local linear quantile regression estimation with jackknife method can be implemented in the R package *quantreg* for each data point U_t by setting covariates as X_i and $X_i(U_i - U_t)$ and the weight as $K_h(U_i - U_t)$, where all data set except X_t are used. Other packages (Matlab or SAS *quantreg* procedure) can be also modified slightly for our estimation method. Obviously, these methods are also applicable to a general local polynomial quantile regression estimation with some necessary modification.

Next, I derive the Bahadur representation of the estimator by using jackknife method and local linear fitting technique. Assume $A_0(U_t)$ is the true coefficient. Define $\varepsilon_t = \psi_\tau(Y_t - A(U_t)^T X_t)$, $\Omega^*(U_t) = E(X_t X_t^T f_{y|u,x}(q_\tau(U_t, X_t))|U_t)$, $\theta = \sqrt{nh}(A(U_t) - A_0(U_t))$, $R_1(U_t) = \frac{1}{n f_u(U_t)}(\Omega^*(U_t))^{-1} \sum_{i \neq t} \varepsilon_i X_i K_h(U_i - U_t)$, and

$$R_2(U_t) = \frac{1}{n f_u(U_t)}(\Omega^*(U_t))^{-1} \sum_{i \neq t}^{n} (\psi_\tau(Y_i^*) - \varepsilon_i) X_i K_h(U_i - U_t),$$

where $\psi_\tau = \tau - I_{\{x<0\}}$.

Theorem 2.1: (Bahadur Representation): Under Assumption A, the local linear

estimator of $A(U_t)$ has the following representation:

$$\hat{\theta} = \frac{(\Omega^*(U_t))^{-1}}{\sqrt{nh}f_u(U_t)} \sum_{i \neq t} \varphi_\tau(Y_i^*) X_i K(\frac{U_i - U_t}{h}) + o_p(1). \tag{2.7}$$

Therefore,

$$\hat{A}_{-t}(U_t) - A(U_t) = R_1(U_t) + R_2(U_t) + o_p(\frac{1}{\sqrt{nh}}), \tag{2.8}$$

where $Y_i^* = Y_i - X_i'(A(U_t) + A^{(1)}(U_t)(U_i - U_t))$.

Proof: See Section 2.8.□

Remark 2.2: The local linear estimator $\hat{A}_{-t}$ is consistent with the optimal nonparametric convergence rate $\sqrt{nh}$.

2.3 Test Statistics and Asymptotic Distribution

2.3.1 Test of Functional Form of Varying Coefficients

Section is devoted to fitting a varying coefficient quantile regression model. Now, it turns to one general and interesting testing problem to check whether the varying coefficient are of some specific functional form. This is equivalent to the testing hypothesis in (2.1). Then, the corresponding generalized quasi-likelihood ratio (GQLR) test statistic is defined as follows

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{j=0}^{p} \hat{a}_{k,-t}(U_t) X_{k,t} \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{k=0}^{p} a_{0,k}(U_t) X_{k,t} \right),
\end{aligned}
$$

where $\ell(H_a) = \sum_{t=1}^{n} \rho_\tau(Y_t - \sum_{k=0}^{p} \hat{a}_{k,-t}(U_t) X_{k,t})$, and $\hat{a}_{k,-t}(U_t)$ is the nonparametric estimate of $a_k(U_t)$ by using local linear estimation technique with jackknife method

under the alternative hypothesis, and $\ell(H_0) = \sum_{t=1}^{n} \rho(Y_t - \sum_{k=0}^{p} a_{k,0}(U_t)X_{k,t})$ with $a_{k,0}(U_t)$ is the true function under the null hypothesis.

To derive the asymptotic distribution of test statistic under H_0, we need the following assumptions in addition to Assumption A.

Assumption B:

(B1) The kernel function $K(\cdot)$ is a symmetric probability density function with bounded support and is Lipschitz continuous.

(B2) $E|\varepsilon|^4 \leq \infty$.

Let

$$
\begin{aligned}
\mu_n &= \frac{p\tau(1-\tau)}{2h} E\left(\frac{1}{f_{y|u,x}(q_\tau(U_t, X_t))}\right) \int K^2(t)dt, \\
\sigma_n^2 &= \frac{2(\tau(1-\tau))^2 p}{h} E\left(\frac{1}{f_{y|u,x}^2(q_\tau(U_t, X_t))}\right) \int ((2K(t) - K * K(t))^2 dt, \\
T_2 &= \sum_{t=1}^{n} \varepsilon_t X_t \frac{h^2}{2} a^{(2)}(U_t)\mu_2 + o(nh^2), \\
T_4 &= \frac{h^2\mu_2}{n} \sum_{t=1}^{n} \frac{f_{y|u,x}(q_\tau(U_t, X_t))a^{(2)}(U_t)}{f_u(U_t)\Omega^*(U_t)} \sum_{j \neq t} \varepsilon_j X_j K_h\left(\frac{U_j - U_t}{h}\right) = O(h^2), \\
T_5 &= \frac{1}{8} \sum_{t=1}^{n} f_{y|u,x}(q_\tau(U_t, X_t) X_t X_t^T h^4 (a^{(2)}(U_t)\mu_2)^2 + o(nh^4),
\end{aligned}
$$

and

$$
d_n = T_2 - T_4 - T_5.
$$

Then, we have the following theorem.

Theorem 2.2: Suppose Assumptions A and B hold, then under H_0 as $h \to 0$ and $nh \to \infty$,

$$
\sigma_n^{-1}(T_n - d_n - \mu_n) \to N(0, 1).
$$

Proof: See Section 2.8.$\square$

Remark 2.3: In general, the above testing approach can be extended to the composite null hypothesis testing hypothesis:

$$H_0 : A_\tau(\cdot) \in \mathscr{A}_{0,\tau} \quad \text{versus} \quad H_a : A_\tau(\cdot) \notin \mathscr{A}_{0,\tau},$$

where $\mathscr{A}_{0,\tau}$ is a set of functions for the given τ. The quasi-likelihood $\mathscr{L}(H_a)$ can be constructed by using local linear estimator and one can get the estimate by either parametric or nonparametric method to build the quasi-likelihood $\mathscr{L}(H_0)$. Let $A_{0,\tau}(\cdot)$ denote the true value of the function $A_\tau(\cdot)$. Then the generalized quasi-likelihood ratio $T_n(A_\theta)$ can be decomposed as

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_{-t}(U_t)^T X_t \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_0^T X_t \right) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_{-t}(U_t)^T X_t \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_0(U_t)^T X_t \right) \\
&\quad + \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_0(U_t)^T X_t \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_0^T X_t \right) \\
&= T_n(A_\theta) - T_n^*(A_\theta),
\end{aligned}
$$

where $T_n(A_\theta)$ is the generalized quasi-likelihood ration for the testing problem

$$H_0 : A(u) = A_0(u) \quad \text{versus} \quad H_a : A(u) \neq A_0(u)$$

and $T_n^*(A_\theta)$ is the generalized quasi-likelihood ratio for the testing problem

$$H_0 : A(u) = A_0(u) \quad \text{versus} \quad H_a : A(u) \in \mathscr{A}_0.$$

Since we do not know the true value $A_0(u)$, the generalized quasi-likelihood ratio for the composite null hypothesis can be decomposed into two generalized quasi-likelihood ratios for two fabricated simple null hypothesis problems. Then, the Bahadur representation and asymptotic distribution can be easily obtained.

2.3.2 Test of Constancy of Varying Coefficient

One special case of the hypothesis in (2.1) is to check whether the coefficient functions are actually varying. This is equivalent to considering the hypothesis testing problem in (2.2) with a known constant vector. By the discussion above, the generalized quasi-likelihood ratio (GQLR) test statistic is defined as follows

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{j=0}^{p} \hat{a}_{k,-t}(U_t) X_{k,t} \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{k=0}^{p} a_{k,0} X_{k,t} \right),
\end{aligned}
$$

where $\ell(H_a) = \sum_{t=1}^{n} \rho_\tau(Y_t - \sum_{k=0}^{p} \hat{a}_{k,-t}(U_t) X_{k,t})$, and $\hat{a}_{k,-t}(U_t)$ is the nonparametric estimate of $a_k(U_t)$ by using local linear estimation with jackknife method under the alternative hypothesis and $\ell(H_0) = \sum_{t=1}^{n} \rho(Y_t - \sum_{k=0}^{p} a_{k,0} X_{k,t})$ with $a_{k,0}$ is the true value under the null hypothesis. With the same notation in Section 2.3.1, we have the following asymptotic result.

Theorem 2.2: Suppose Assumptions A and B hold, then under H_0, as $h \to 0$ and $nh \to \infty$, we have

$$
\sigma_n^{-1}(T_n - \mu_n) \to N(0, 1).
$$

Proof: See Section 2.8.$\square$

2.3.3 Test of Constancy of Varying Coefficient with Unknown Value

In some applications, it may be more interesting in checking the constancy of the varying coefficient with the true value A_0 unknown. Therefore, we consider the

test statistic for the hypothesis in (2.2) with a unknown constant vector. Under the null hypothesis, one can estimate the coefficient $\hat{a}_{0,k}$ for the linear quantile regression and construct the quasi-likelihood as follows

$$\ell(H_0) = \sum_{t=1}^{n} \rho(Y_t - \sum_{k=0}^{p} \hat{a}_{0,k} X_{k,t}).$$

Then, the generalized quasi-likelihood ratio (GQLR) test statistic for hypothesis testing problem in (2.2) with unknown A_0 is defined by

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{j=0}^{p} \hat{a}_{k,-t}(U_t) X_{k,t} \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{k=0}^{p} \hat{a}_{0,k} X_{k,t} \right) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{j=0}^{p} \hat{a}_{k,-t}(U_t) X_{k,t} \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{k=0}^{p} a_{0,k} X_{k,t} \right) \\
&\quad + \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{k=0}^{p} a_{0,k} X_{k,t} \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \sum_{k=0}^{p} \hat{a}_{0,k} X_{k,t} \right) \\
&\equiv T_{n1} + T_{n2}.
\end{aligned}
$$

Remark 2.4: As shown in Koenker (2005), the limiting distribution of T_{n2} is a χ^2 distribution for a parametric quantile regression model.

With the same notation in Section 2.3.1, we have the following asymptotic result.

Theorem 2.4: Suppose Assumptions A and B hold, then under H_0, as $h \to 0$ and $nh \to \infty$, we have

$$\sigma_n^{-1}(T_n - \mu_n) \to N(0,1).$$

Proof: See Section 2.8.□

2.4 Estimation of Covariance Matrix and Bandwidth Selection

In Theorems 2.2 - 2.4, the variance of the limit distribution involves the conditional density of Y, $f_{y|u,x}(q_\tau(U_t, X_t))$. For the purpose of statistical inference, we need to obtain a consistent estimate for $f_{y|u,x}(q_\tau(U_t, X_t))$. As pointed out by Cai and Xu (2008), there are two methods are available for the consistent estimate. The first one is the difference quotients method of Koenker and Xiao (2004), such that

$$\hat{f}_{y|u,x}(q_\tau(u,x)) = \frac{\tau_j - \tau_{j-1}}{q_{\tau_j}(u,x) - q_{\tau_{j-1}}(u,x)}$$

for some appropriately chosen sequence of $\{\tau_j\}$. $q_\tau(\cdot)$ can be estimated by using a variant of the empirical quantile function for the linear model $\hat{q}(\tau|x) = x^T \hat{\alpha}(\tau)$ proposed in Bassett and Koenker (1982). Then, the conditional density can be estimated by

$$\hat{f}_{y|u,x}(q_\tau(u,x)) = \frac{\tau_j - \tau_{j-1}}{x^T(\hat{\alpha}(t+h_n) - \hat{\alpha}(t-h_n))}.$$

The second one is the Nadaraya-Watson type double-kernel method proposed in Fan, Yao, and Tong (1996) and they estimated the conditional density functions and square roots and their partial derivatives directly by using locally polynomial regression. Assume $g(y|x)$ is the conditional density of Y given X. As $h_2 \to 0$,

$$E\{K_{h_2}(Y - y)|X = x\} \approx g(y|x), \tag{2.9}$$

where $K(\cdot)$ is a nonnegative density function. It can be considered as regression of $K_{h_2}(Y - y)$ on X. By using the locally quadratic fitting technique one can estimate the coefficient by minimize

$$\sum_{i=1}^n [K_{h_2}(Y_i - y) - \beta_0 - \beta_1^T(X_i - x) - \beta_2^T vec(X_i - x)(X_i - x)^T]^2 W_{h_1}(X_i - x),$$

where $W(\cdot)$ is a nonnegative kernel function and h_1 is the bandwidth. One can estimate $\hat{g}(y|x) = \hat{\beta}_0$. Here

$$\hat{\beta} := (\hat{\beta}_0, \hat{\beta}_1^T, \hat{\beta}_2^T)^T = (\mathscr{X}^T \mathscr{W} \mathscr{X})^{-1} \mathscr{X}^T \mathscr{W} \mathscr{Y},$$

where $\mathscr{X}$ is the design matrix of the above least squares problem, $\mathscr{W} = diag(W_{h_1}(X_1 - x), ..., W_{h_1}(X_n - x))$, and $\mathscr{Y} = (K_{h_2}(Y_1 - y), ..., K_{h_2}(Y_n - y))^T$. For the univariate x, we have

$$\hat{g}(y|x) = h_1^{-1} \sum_{i=1}^{n} W_0^n(\frac{X_i - x}{h_1}) K_{h_2}(Y_i - y),$$

where $W_0^n(t) = \tau_0^T S_n^{-1}(1, h_1 t, h_1^2 t^2)^T \times W(t)$ with τ_0 the unit vector with 1st element 1, and

$$S_n = \begin{pmatrix} s_{n,0} & s_{n,1} & s_{n,2} \\ s_{n,1} & s_{n,2} & s_{n,3} \\ s_{n,2} & s_{n,3} & s_{n,4} \end{pmatrix},$$

where $s_{n,j} = \sum_{i=1}^{n}(X_i - x)^j W_{h_1}(X_i - x)$. Fan, Yao, and Tong (1996) also proposed to select the optional bandwidth h_1 for estimating $g'(y|x)$, which is given by

$$\hat{h}_1(y) = c_0 * \text{argmin}_h \int RSC(x, y; h) dx, \tag{2.10}$$

where $RSC(x, y; h_1) = \sigma^2(x, y; h_1)(1 + 3V_n(x; h_1))$, $V_n(x, h_1)$ is the first diagonal element of matrix $S_n^{-1} T_n S_n^{-1}$, and c_0 is a positive constant. As for the bandwidth h_2, they adopted

$$\hat{h}_2 = \left[\frac{8\sqrt{\pi}\nu_0}{3 \mu_2^2} \right]^{1/5} s_y n^{-1/5}, \tag{2.11}$$

where s_y is the sample standard deviation of Y and μ_2 and ν_0 are defined in Theorem 1.2.

As indicated in Chen and Linton (2001), for the conditional density function

$f(y|x) = \frac{f(y,x)}{f(x)}$, one can estimate $f(y|x)$ by

$$\hat{f}_{h_1h_2h_3}(y|x) = \frac{\hat{f}_{h_1h_2}(y,x)}{\hat{f}_{h_3}(x)},$$

where

$$\hat{f}_{h_1h_2}(y,x) = \frac{1}{nh_1h_2}\sum_{i=1}^{n}K(\frac{y-Y_i}{h_1})K(\frac{x-X_i}{h_2}),$$

$$\hat{f}_{h_3}(x) = \frac{1}{nh_3}\sum_{i=1}^{n}K(\frac{z-Z_i}{h_3}),$$

$K(\cdot)$ is a bounded kernel function, and h_1, h_2 and h_3 are positive bandwidth sequences, which decay to zero as $n \to \infty$. By fixing $h_1 = h_2$, they derived the following asymptotic result.

Theorem 2.5: (Theorem 3 in Chen and Linton (2001)) Under some assumptions and let $h_1, h_3 > 0$ be such that

$$\lim_{n\to\infty}[n \times \min(h_1^2, h_3) \times \max(h_1^4, h_3^4)] = 0$$

$$\lim_{n\to\infty}\min(1, \frac{h_3}{h_1^2}) + lim_{n\to\infty}\min(1, \frac{h_1^2}{h_3}) > 0 \quad \text{exists.} \tag{2.12}$$

Then,

$$\sqrt{n \times \min(h_1^2, h_3)}(\hat{f}_{h_1h_2h_3}(y|z) - f(y|z)) \to N(0, V(y,z)),$$

where

$$V(y,z) = \frac{f(y|z)}{f(z)}\nu_0[\min(1, \lim_n \frac{h_3}{h_1^2}\nu_0) + \min(1, \lim_n \frac{h_1^2}{h_3})f(y|z)].$$

Hence, with $nh_1^2 \to \infty$, and $nh_1^6 \to 0$, for case $h_1 = h_2 = \sqrt{h_3}$, we have

$$\sqrt{nh_1^2}(\hat{f}_{h_1h_2h_3}(y|z) - f(y|z)) \to N(0, V_1(y,z) + V_2(y,z))$$

and for case $h_1 = h_2 = h_3$,

$$\sqrt{nh_1^2}(\hat{f}_{h_1 h_2 h_3}(y|z) - f(y|z)) \to N(0, V_2(y, z)),$$

where $V_1(y, \mu) = \mu_2 \sqrt{2\pi}\sigma(1 - \rho^2)\exp\{-\frac{(y-\mu)^2}{\sigma^2(1-\rho^2)}\}$, and $V_2(y, \mu) = \nu_0^2(2\pi)^{1/2}\sigma$.

From the above results, I will use the estimator for the conditional density as follows:

$$\hat{f}_{y|u,x}(q_\tau(u, x)) = \frac{\sum_{t=1}^n K_{h_1}(U_t - u, X_t - x)L_{h1}(Y_t - q_\tau(u, x))}{\sum_{t=1}^n K_{h_2}(U_t - u, X_t - x)},$$

where $L(\cdot)$ and $K(\cdot)$ are kernel functions and the bandwidths satisfy $h_1 = \sqrt{h_2}$ and $nh_1^2 \to \infty, nh_1^6 \to 0$.

Remark 2.5: One can prove that this estimator is the consistent estimator of $f_{y|u,x}(q_\tau(u, x))$ by using the theoretical result of Theorem 2.5.

Set

$$
\begin{aligned}
\hat{\mu}_n &= \frac{p\tau(1-\tau)}{2h}E\Big(\frac{1}{\hat{f}_{y|u,x}(q_\tau(U_t, X_t))}\Big)\nu_0, \\
\hat{\sigma}_n^2 &= \frac{2(\tau(1-\tau))^2 p}{h}E\Big(\frac{1}{\hat{f}_{y|u,x}^2(q_\tau(U_t, X_t))}\Big)\int ((2K(t) - K * K(t))^2 dt,
\end{aligned}
$$

and

$$\hat{d}_n = T_2 - \hat{T}_4 - \hat{T}_5.$$

Then, we have the following theorem.

Theorem 2.6: Suppose Assumptions A and B hold, $h_1 = \sqrt{h_2}$, $nh_1^6 \to 0$, and $nh_1^2 \to \infty$. Then, under H_0 in (2.2), we have

$$\hat{\sigma}_n^{-1}(T_n - \hat{d}_n - \hat{\mu}_n) \to N(0, 1).$$

Proof: See Section 2.8.□

Remark 2.6: We can get the similar asymptotic results as in Theorems 2.3 and 2.4 and show that the estimator for the test statistic is consistent. Therefore, we can use simulation method to approximate the distribution of the proposed test statistic.

2.5 Power of Test

In this section, we consider the power of the quasi-likelihood ratio test based on local linear fit. For simplicity, we only focus on the null hypothesis in (2.2) with a known vector. The power of the test is considered under local alternatives as follows

$$H_a : A(u) = A_0(u) + \frac{1}{\sqrt{nh}}\Delta(u),$$

where $\Delta(u) = (\Delta_1(u), \Delta_2(u), ..., \Delta_p(u))^T$ is a vector of functions, satisfying $E(||\Delta(u)||_2^2) < \infty$ and $A_0(u)$ is a known constant under H_0. Define

$$\sigma_n^* = \sqrt{\sigma_n^2 + (\tau(1-\tau))E\left[\Delta^T(U)XX^T\Delta(U)\right]}$$

and $d_{2n} = \frac{1}{2h}E(f_{y|u,x}(q_\tau|X_t, U_t)\Delta(U_t)^T X_t X_t^T \Delta(U_t))$, which goes to infinity when $h \to 0$ if $\Delta(\cdot)$ is nonzero. Then, we have the asymptotical distribution of test statistic T_n.

Theorem 2.7: Suppose Assumptions A and B hold,and assume A_0 is true constant coefficient. By using the similar notation in Theorem 2.2, then under H_a, we have

$$\sigma_n^{*-1}(T_n - \mu_n + d_{2n}) \to N(0, 1).$$

Proof: See Appendix.□

Remark 2.7: From Theorem 2.7, it is easy to see that the test statistic diverges if $\Delta(\cdot)$ departs from zero. This implies that the test is consistent.

2.6   Empirical Examples

In this section, we conduct three Monte Carlo simulated examples to examine the finite sample performance of the proposed test procedure. In our computation, the Gaussian Kernel $K(u) = \frac{1}{\sqrt{2\pi}}e^{\frac{-u^2}{2}}$ is used. Simulation procedures and results are reported next.

2.6.1   Simulation Procedures

Instead of using the limiting distribution to compute the critical values of the proposed test statistics, we suggest using the simulation approach, which might give a better finite sample performance. The simulation procedure is briefly described as follows:

1) Under H_0, we estimate the unknown coefficients in model $q_\tau(U_t, X_t) = A_{0,\tau}^T X_t$ at the τth quantile using the parametric method and obtain the sum $\ell(H_0) = \sum_{t=1}^n \rho_\tau(y_t - \hat{A}_{0,\tau}^T X_t)$, which becomes to $\ell(H_0) = \sum_{t=1}^n \rho_\tau(y_t - A_{0,\tau}(U_t)^T X_t)$ if $A_{0,\tau}(u)$ in known under H_0.

2) Under H_a, we estimate the coefficients in model $q_\tau(U_t, X_t) = A_\tau(U_t)^T X_t$ at the same quantile τ using the jackknife method and local fitting technique and obtain the sum $\ell(H_a) = \sum \rho_\tau(y_t - \hat{A}_{-t,\tau}(U_t)^T X_t)$.

3) Calculate the test statistic $T_n = \ell(H_a) - \ell(H_0)$.

4) Repeat Steps (1) - (3) a large number of times, say 1000 times, to find the empirical distribution of $\{T_n\}$. The critical value at significance level α is given by the $(1 - \alpha)th$ quantile.

Remark 2.8: From Theorems 2.2 - 2.4, the consistent estimators for μ_n and σ_n^2 are provided, so that the critical value can be obtained from the above simulation procedure.

Remark 2.9: In Procedure 1), the estimation is not necessary for hypothesis testing problems in (2.1)with a known functional vector and (2.2) with a known vector since

the true values are given in the null hypothesis.

2.6.2   Simulation Results

Example 2.1: In this simulated example, I consider the following data generating process.

$$Y_t = a_1(U_t)X_t + e_t, \quad 1 \leq t \leq n, \tag{2.13}$$

where $a_1(u) = 2$, U_t is generated from uniform $(0,3)$ independently, $e_t \sim N(0, 0.3)$ and $X_t \sim N(0.5, 0.4)$. Then, the corresponding quantile regression is

$$q_\tau(U_t, X_t) = a_{0,\tau}(U_t) + a_{1,\tau}(U_t)X_t,$$

where $a_{0,\tau}(U_t) = \sqrt{0.3}\Phi^{-1}(\tau)$, $a_{1,\tau}(u) = a_1(u)$, and $\Phi^{-1}(\tau)$ is the τ-th quantile of $N(0,1)$.

In this example, I consider $a_1(u) = 2$ as the known coefficient under the null hypothesis. I choose the sample sizes as $n = 250$, 500 and 800 and repeat the simulation $m = 1000$ times. I report the simulation results for the testing nominal sizes at 1%, 5% and 10% for different quantiles $\tau = 0.2$, 0.4, 0.6 and 0.8 in Table 2.1, from which, we can see the empirical sizes of the proposed test statistic at different significance levels and different quantiles are close the true nominal sizes. This implies that the proposed test can deliver a correct test size.

To demonstrate the power of the proposed test, the power function is evaluated under a sequence of the alternative models indexed by λ

$$H_a : a_{1,\tau}(u) = 2 + \frac{\lambda}{\sqrt{nh}}\Delta(u), \quad 0 \leq \lambda \leq 3, \tag{2.14}$$

where $\Delta(u) = 2\sin(\sqrt{2}\pi u)$ and $\lambda_i = 0.2\,i$ for $1 \leq i \leq 15$. The simulation is repeated 1000 times for each sample size $n = 250$, $n = 500$ and $n = 800$ and for each quantile $\tau = 0.2$, 0.4, 0.6, and 0.8. Given the significance level 5%, the power function $p(\lambda)$

Table 2.1: : Finite sample rejection rates for example 2.1

nominal size	sample size	$\tau = 0.2$	$\tau = 0.4$	$\tau = 0.6$	$\tau = 0.8$
	$n = 250$	0.099	0.122	0.109	0.125
10%	$n = 500$	0.093	0.080	0.115	0.119
	$n = 800$	0.107	0.093	0.119	0.109
	$n = 250$	0.054	0.068	0.058	0.050
5%	$n = 500$	0.047	0.069	0.072	0.071
	$n = 800$	0.060	0.052	0.061	0.053
	$n = 250$	0.020	0.024	0.023	0.013
1%	$n = 500$	0.021	0.018	0.025	0.014
	$n = 800$	0.015	0.016	0.017	0.019

is estimated based on the relative frequency of T_n over the 1000 simulations. I plot the power curves in Figure 2.1 for all settings. One can see from Figure 2.1 that the power curves are almost same for three sample sizes. This observation is consistent with our local alternative setting. Also, one can observe that indeed, the proposed test statistic is powerful.

Example 2.2: In this simulated example, the data generating process and the settings are exactly the same as those in Example 2.1. But under the null hypothesis, A_0 is unknown and it needs an estimation. Different from Example 2.1, I adopt the linear quantile regression to estimate this coefficient under the null hypothesis. The simulated sizes for all settings are listed in Table 2.2 and we plot the power curves in

Table 2.2: : Finite sample rejection rates for example 2.2

nominal size	sample size	$\tau = 0.2$	$\tau = 0.4$	$\tau = 0.6$	$\tau = 0.8$
	$n = 250$	0.107	0.121	0.103	0.111
10%	$n = 500$	0.121	0.121	0.114	0.111
	$n = 800$	0.084	0.115	0.113	0.104
	$n = 250$	0.062	0.061	0.067	0.058
5%	$n = 500$	0.060	0.061	0.061	0.064
	$n = 800$	0.055	0.061	0.068	0.053
	$n = 250$	0.019	0.017	0.021	0.019
1%	$n = 500$	0.022	0.018	0.018	0.015
	$n = 800$	0.014	0.014	0.024	0.018

Figure 2.2. From both Table 2.2 and Figure 2.2, the same conclusions as in Example 2.1 can be made.

Example 2.3: The purpose of this example is to test whether the coefficient functions have a specific known form. To this end, I consider the simplest model

$$Y_t = a_1(U_t)X_{1t} + a_2(U_t)X_{2t} + e_t, \quad 1 \le t \le n, \tag{2.15}$$

where $a_1(u) = \sin(\sqrt{2}\pi u)$, $a_2(u) = \cos(\sqrt{2}\pi u)$, U_t is generated from uniform $(0,3)$ independently, $e_t \sim N(0,0.3)$, $X_{1t} \sim N(0.5,0.4)$ and $X_{2t} \sim N(0.75,0.4)$. Then, the corresponding quantile regression is

$$q_\tau(U_t, X_t) = a_{0,\tau}(U_t) + a_{1,\tau}(U_t)X_{1t} + a_{2,\tau}(U_t)X_{2t},$$

where $a_{0,\tau}(u) = \sqrt{0.3}\Phi^{-1}(\tau)$, $a_{1,\tau}(u) = a_1(u)$, $a_{2,\tau}(u) = a_2(u)$, and $\Phi^{-1}(\tau)$ is the τ-th quantile of the $N(0,1)$. The remaining settings are the same as those in Example 2.1. The simulated sizes for all settings are listed in Table 2.3. Similar to Examples 2.1 and 2.2, the same observations can be made. Similarly, the power function is

Table 2.3: : Finite sample rejection rates for example 2.3

nominal size	sample size	$\tau = 0.2$	$\tau = 0.4$	$\tau = 0.6$	$\tau = 0.8$
	$n = 250$	0.094	0.111	0.111	0.104
10%	$n = 500$	0.116	0.109	0.098	0.108
	$n = 800$	0.098	0.112	0.106	0.106
	$n = 250$	0.042	0.063	0.061	0.053
5%	$n = 500$	0.055	0.051	0.048	0.060
	$n = 800$	0.052	0.055	0.061	0.061
	$n = 250$	0.012	0.025	0.015	0.018
1%	$n = 500$	0.017	0.016	0.015	0.017
	$n = 800$	0.017	0.016	0.022	0.015

evaluated under a sequence of the alternative models indexed by $0 \leq \lambda \leq 1$,

$$H_a : a_{1,\tau}(u) = \sin(\sqrt{2}\pi\, u) + \frac{\lambda}{\sqrt{nh}}\Delta(u) \quad \text{and} \quad a_2(u) = \cos(\sqrt{2}\pi u) + \frac{\lambda}{\sqrt{nh}}\Delta(u),$$

where $\Delta(u) = u^4 e^{-u/10}$ and $\lambda = 0.05\, i$ for $1 \leq i \leq 20$. Given the significance level 5%, we compute the power curves as functions of λ and plot them in Figure 2.3, from which we can conclude that the proposed test statistic indeed is powerful.

2.7 A Real Example

In previous section, I conducted Monte Carlo simulation to illustrate the effectiveness and the validity of the proposed test statistics. In this section, I consider the application of these methodologies to a real example.

Here I analyze a subset of the Boston house price data (http://lib.stat.cum.edu /datasets/boston) of Harrison and Rubinfeld (1978) which is used to study the effect of air pollution on real estate price in the greater Boston area in 1970s. The data set consist of 506 observations on 14 variables. For the complete description of all there variables, the reader is referred to the papers by Harrison and Rubinfeld (1978) and Gilley and Pace (1996). As indicated in Cai and Xu (2008) which analyzed this data set by using a varying coefficient quantile regression model, we focus on exploring the possible (linear, nonparametric or semiparametric) relationships between the dependent variable and some major factors which might affect the house price. Here I adopt the same notation as in Cai and Xu (2008) in order to do a comparison.

Y: the dependent variable, the median value of owner-occupied homes in \$1,000's (house price).

U: proportion of population of lower educational status.

X_1: the average number of rooms per house in the area.

X_2: the per capital crime rate by town.

X_3: the full property tax rate per \$10,000.

X_4: the pupil/teacher ratio by town school district.

Note that there are many papers investigating this data set in the literature, and the reader is referred to the paper by Cai and Xu (2008) for details.

In this section, I will focus on two models. First, we consider the model from Cai and Xu (2008) which is the following quantile smooth coefficient model

$$q_\tau(U_t, X_t) = a_{0,\tau}(U_t) + a_{1,\tau}(U_t)X_{t1} + a_{2,\tau}(U_t)X_{t2}^*, \qquad (2.16)$$

where $X_{t2}^* = \log(X_{t2})$. Our interest is to check whether the functional coefficients in $A_\tau(u) = (a_{0,\tau}(u), a_{1,\tau}(u), a_{2,\tau}(u))^T$ in model (2.16) are indeed varying with u. That is to test the null hypothesis $H_0 : A_\tau(u) = A_0$, where $A_{0,\tau}$ is a vector of unknown parameters. For this testing problem, I calculate the test statistic by using the proposed test procedure. The corresponding p-value are reported in Table 2.4. Therefore, one can see that all the p-values are less than significant level 0.05 from Table 2.4, which implies that the varying coefficients are indeed varying.

Table 2.4: : The p-values for testing constancy in model (2.16)

τ	0.2	0.4	0.6	0.8
p-value	0.000	0.020	0.000	0.000

It is clear that model (3.9) does not include two variables X_3 and X_4. The reason as claimed by Cai and Xu (2008) is that the functional coefficients for variables X_3 and X_4 may be constant. Therefore, I use the proposed test procedure to test whether the coefficients of X_3 and X_4 are constant or not. To this effect, we consider the following model

$$q_\tau(U_t, X_t) = a_{0,\tau}^*(U_t) + a_{3,\tau}(U_t)X_{t3} + a_{4,\tau}(U_t)X_{t4}^*, \qquad (2.17)$$

and then consider the testing problem formulated as the null hypothesis H_0 :

$A_\tau^*(u) = A_0^*$, where $A_\tau^*(u) = (a_{0,\tau}^*(u), a_{3,\tau}(u), a_{4,\tau}(u))^T$ and $A_{0,\tau}^*$ is a vector of unknown parameters. By using the test statistic by following the test procedure as in Section 2.3.3, I calculate the quasi-likelihood using linear parametric quantile regression under the null hypothesis and calculate the quasi-likelihood using local linear fitting method. The corresponding p-values are reported in Table 2.5, from which, one can see that all the p-values are greater than significant level 0.05. This implies that the varying coefficients are indeed constant.

Table 2.5: : The p-values for testing constancy in model (2.17)

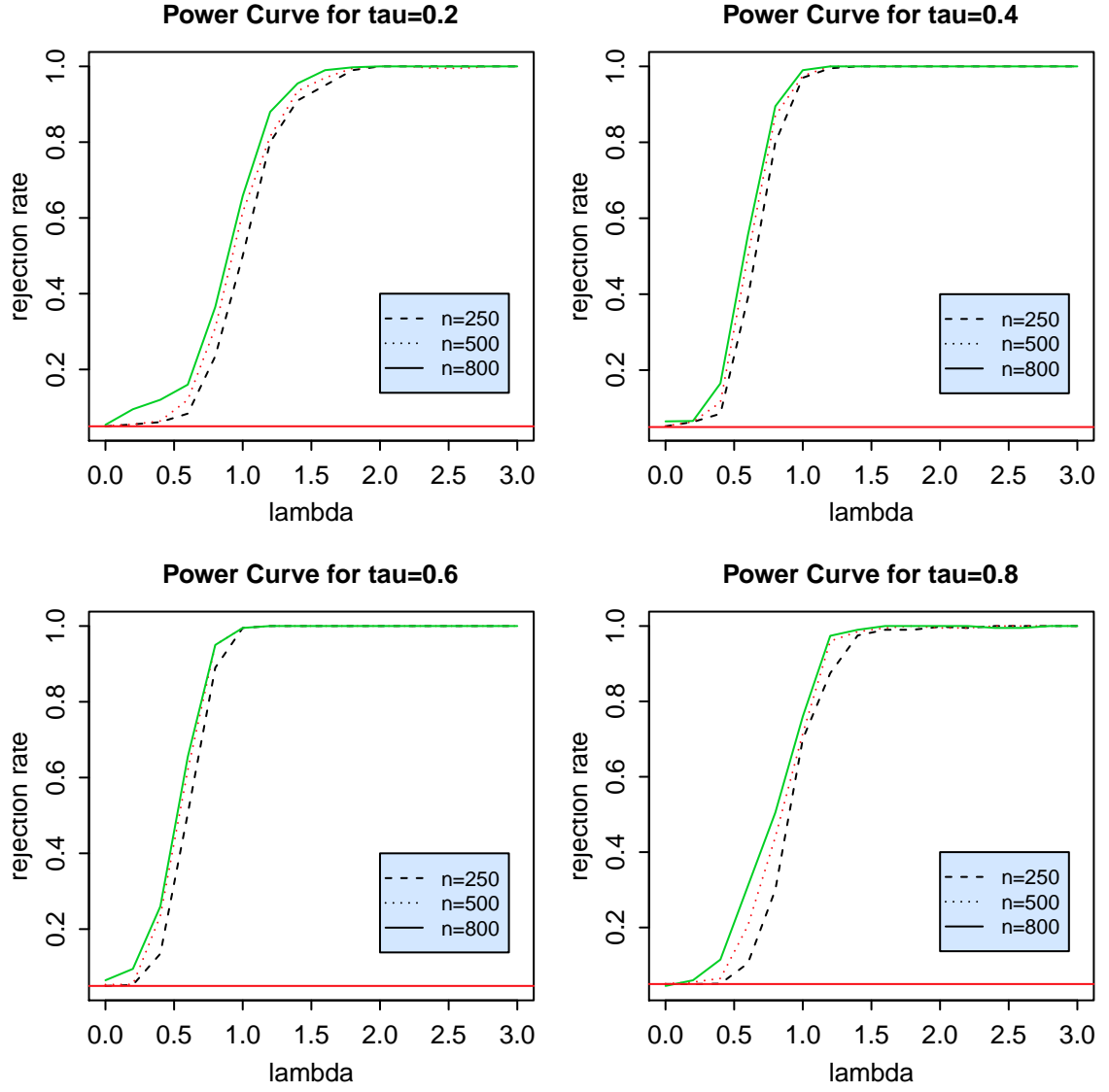| τ	0.2	0.4	0.6	0.8
p-value	0.680	0.910	0.220	0.340

Figure 2.1: The plot of power curves for the testing hypothesis in Example 2.1 with the nominal size 5% in Section 2.6. The dashed line is for $n = 250$, the solid line is for $n = 500$ and the dashed-dotted line is for $n = 800$.
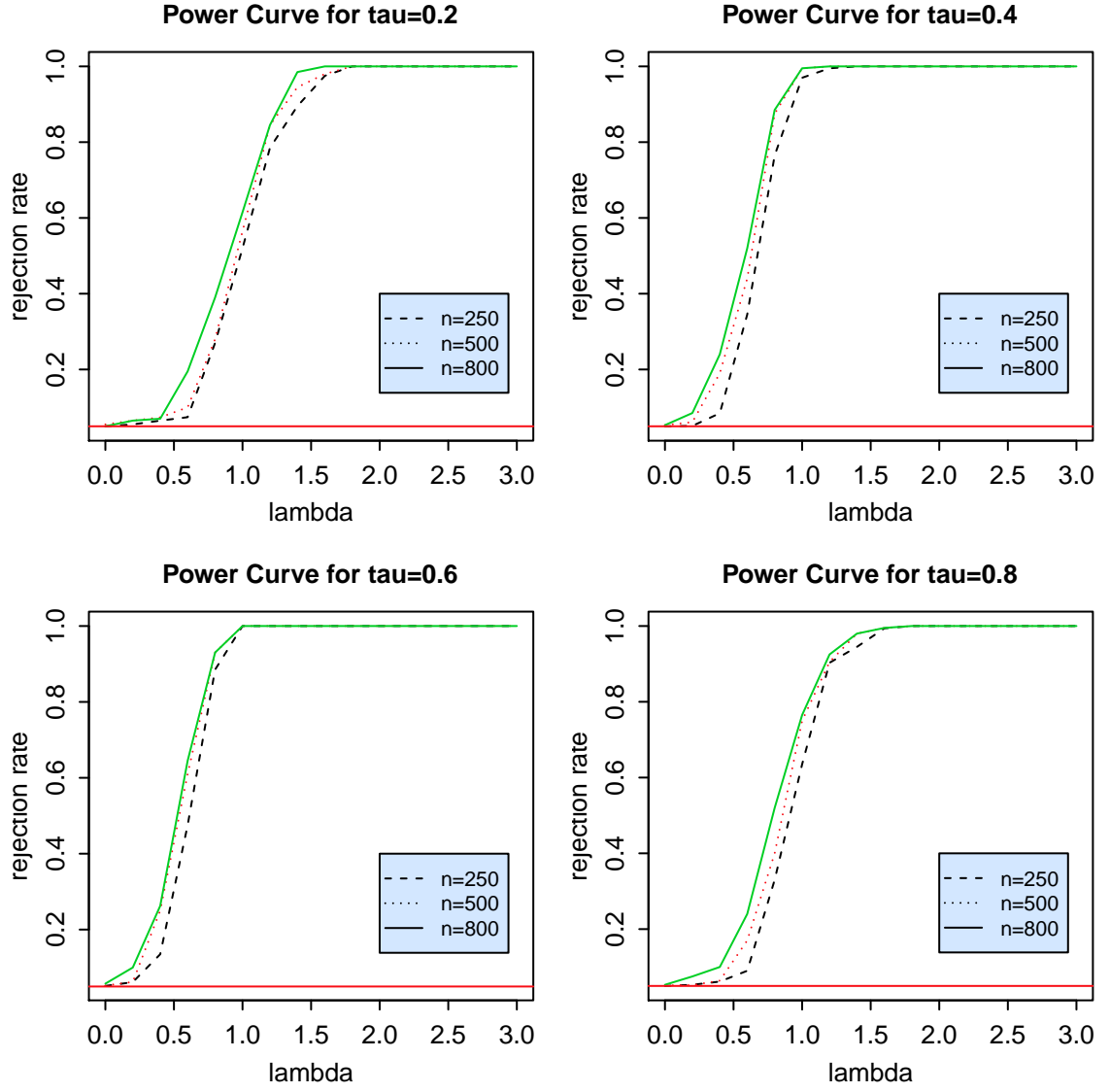
Figure 2.2: The plot of power curves for the testing hypothesis in Example 2.2 with the nominal size 5% in Section 2.6. The dashed line is for $n = 250$, the solid line is for $n = 500$ and the dashed-dotted line is for $n = 800$.
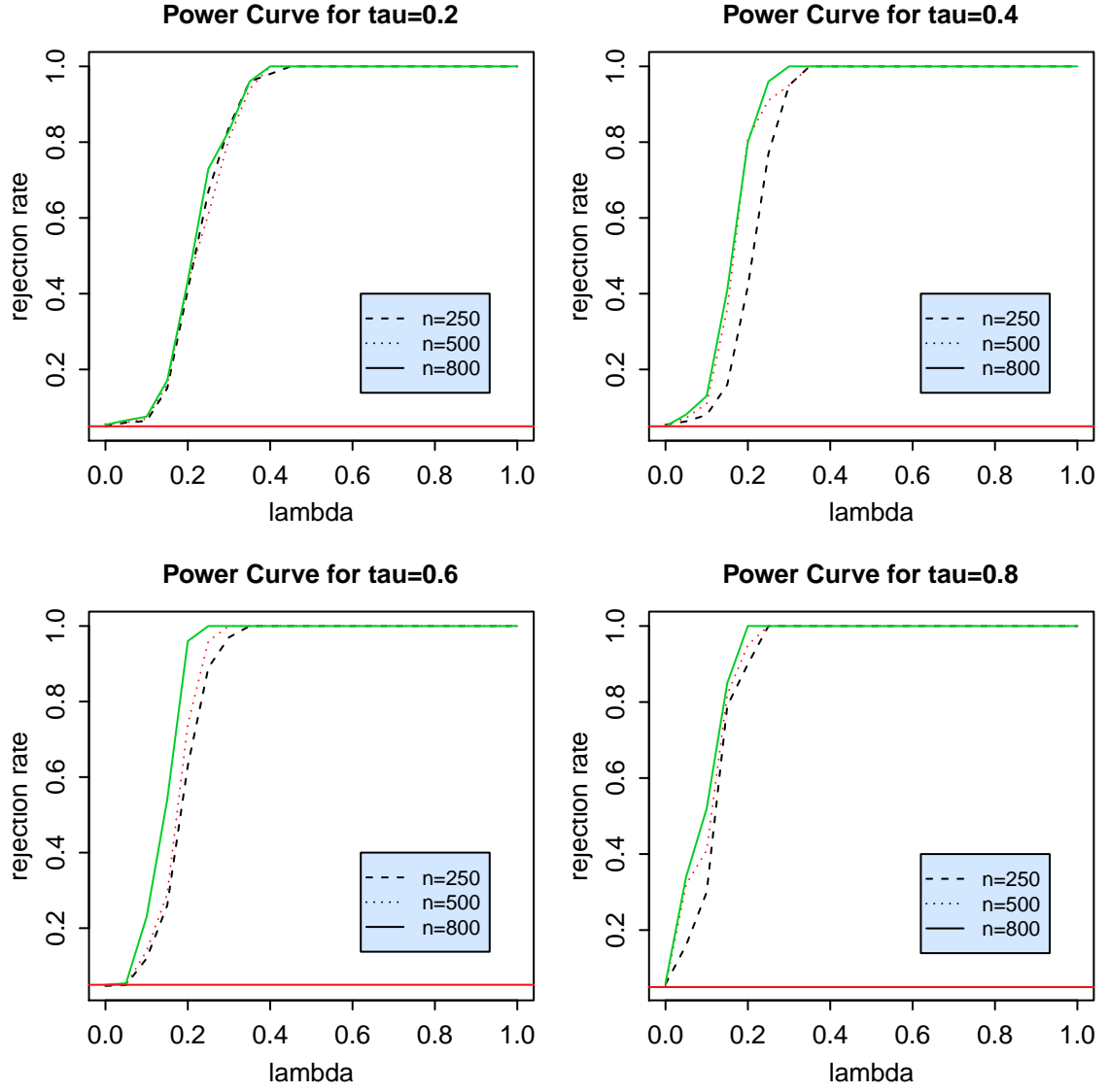
Figure 2.3: The plot of power curves for the testing hypothesis in Example 2.3 with the nominal size 5% in Section 2.6. The dashed line is for $n = 250$, the solid line is for $n = 500$ and the dashed-dotted line is for $n = 800$.

2.8 Complements

In this section, we give the derivations of the main results presented in previous sections of this chapter. Before moving to the detailed proofs, we need the following lemma.

Lemma 1: (Lemma 3.3.2 in Zhang (2000))

(1) Assume a_2, b_2 are finite and $[a_2, b_2] \subseteq [a_1, b_1]$. Suppose $g_1(x)$ is Lipschitz continuous in $[a_1, b_1]$ and $g_2(y)$ is continuous in $[a_2, b_2]$; $K(x)$ is a symmetric function with a bounded support and $\int y^q K^{m_1}(y) dy < \infty$ for some integer $m_1 \geq 1, q \geq 0$. Then as $h \to 0$

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} g_1(x) g_2(y) (\frac{x-y}{h})^q K_h^{m_1}(x-y) dx dy = \frac{\int_{a_2}^{b_2} g_1(x) g_2(x) dx \int y^q K^{m_1}(y) dy + O(h)}{h^{m_1 - 1}}.$$

(2) Assume a_1, b_1 are finite and $[a_1, b_1] \subseteq [a_2, b_2], [a_1, b_1] \subseteq [a_3, b_3]$. Suppose $g_2(y)$ is Lipschitz continuous in $[a_2, b_2]$ and $g_3(z)$ is continuous in $[a_3, b_3]$; $g_1(x)$ is continuous in $[a_1, b_1]$; $K(x)$ is a symmetric function with a bounded support and $\int y^q K^{m_1}(y) dy < \infty$ for some integer $m_1 \geq 1, m_2 \geq 1, q \geq 0$. Then as $h \to 0$

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} g_1(x) g_2(y) g_3(u) (\frac{x-z}{h})^q K_h^{m_1}(x-z) K_h^{m_2}(z-y) dx dy dz$$

$$= \frac{\int_{a_1}^{b_1} g_1(x) g_2(x) g_3(x) dx \int z^q K^{m_1}(z) dz \int K^{m_2}(y) dy + O(h)}{h^{m_1 + m_2 - 2}}.$$

(3) Assume a_1, b_1 are finite and $[a_1, b_1] \subseteq [a_2, b_2], [a_1, b_1] \subseteq [a_3, b_3]$. Suppose $g_2(y)$ is Lipschitz continuous in $[a_2, b_2]$ and $g_3(u)$ is continuous in $[a_3, b_3]$; $g_1(x)$ is continuous in $[a_1, b_1]$; $K(x)$ is a symmetric function with a bounded support and $\int y^q K^{m_1}(y) dy < \infty$ for some integer $m_i \geq 1, q_i \geq 0 (i = 1, 2, 3)$. Then as $h \to 0$

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} g_1(x) g_2(y) g_3(u) (\frac{u-x}{h})^{q_1} K_h^{m_1}(u-x) (\frac{y-x}{h})^{q_2}$$

$$K_h^{m_2}(x-y)(\frac{u-y}{h})^{q_3}K_h^{m_3}(y-u)dxdydu$$

$$=\frac{\int_{a_1}^{b_1}g_1(x)g_2(x)g_3(x)dx\int u^{q_1}K^{m_1}(u)du\int y^{q_2}K^{m_2}(y)(u-y)^{q_3}K^{m_3}(u-y)dy+O(h)}{h^{m_1+m_2+m_3-2}}$$

$$=\frac{\int_{a_1}^{b_1}g_1(x)g_2(x)g_3(x)dx\int u^{q_1}K^{m_1}(u)du\int(y+u)^{q_2}K^{m_2}(y+u)(-y)^{q_3}K^{m_3}(y)dy}{h^{m_1+m_2+m_3-2}}.$$

(4) Assume a_1, b_1 are finite and $[a_1, b_1] \subseteq [a_2, b_2], [a_1, b_1] \subseteq [a_3, b_3], [a_1, b_1] \subseteq [a_4, b_4]$. Suppose $g_2(y)$ is Lipschitz continuous in $[a_2, b_2]$ and $g_3(u)$ is continuous in $[a_3, b_3]$; $g_4(v)$ is continuous in $[a_4, b_4]$; $g_1(x)$ is continuous in $[a_1, b_1]$; $K(x)$ is a symmetric function with a bounded support and $\int y^q K^{m_1}(y)dy < \infty$ for some integer $m_i \geq 1, q_i \geq 0 (i = 1, 2, 3, 4)$. Then as $h \to 0$

$$\int_{a_1}^{b_1}\int_{a_2}^{b_2}\int_{a_3}^{b_3}\int_{a_4}^{b_4}g_1(x)g_2(y)g_3(u)g_4(v)(\frac{u-x}{h})^{q_1}K_h^{m_1}(u-x)(\frac{v-x}{h})^{q_2}K_h^{m_2}(v-x)$$

$$(\frac{u-y}{h})^{q_3}K_h^{m_3}(u-y)(\frac{v-y}{h})^{q_4}K_h^{m_4}(v-y)dxdydudv$$

$$=\frac{\int_{a_1}^{b_1}g_1(x)g_2(x)g_3(x)g_4(x)dx}{h^{m_1+m_2+m_3+m_4-3}}\times$$

$$\frac{\int_{2B}^{2B}[\int_B^B u^{q_1}K^{m_1}(u)(u-y)^{q_3}K^{m_3}(y-u)]du[\int_B^B v^{q_2}K^{m_2}(v)(v-y)^{q_4}K^{m_4}(y-v)]dv]dy}{h^{m_1+m_2+m_3+m_4-3}}$$

$$+\frac{O(h)}{h^{m_1+m_2+m_3+m_4-3}}.$$

Definition 2: (Definition 1 in de Jong (1987)) $W(n)$ is called clean if the conditional expectations of W_{ij} vanish:

$$E(W_{ij}|X_i) = 0 \quad a.s. \quad \text{for all} \quad i, j \leq n.$$

Lemma 2: (Theorem2.1 in de Jong (1987))

Let $W(n)$ be clean with variance $\sigma^2(n)$. Assume

    a) $\sigma(n)^{-2} max_{1 \le i \le n} \sum_{j=1}^{n} \sigma_{ij}^2 \to 0$, as $n \to \infty$.

    b) $\sigma(n)^{-4} E(W(n)^4) \to 3$ as $n \to \infty$.

then

$$\sigma(n)^{-1} W(n) \to N(0,1), \text{as} \quad n \to \infty.$$

Lemma 3: (Proposition 3.2 in de Jong (1987))

Let $W(n)$ be clean and let G_I, G_{II} and G_{IV} be of lower order than $\sigma^4(n)$, then,

$$\sigma(n)^{-1} W(n) \to N(0,1), \text{as} \quad n \to \infty.$$

where

$$
\begin{aligned}
G_I &= \sum_{1 \le i < j \le n} E(W_{ij}^4), \\
G_{II} &= \sum_{1 \le i < j \le n} \{E(W_{ij}^2 W_{ik}^2) + E(W_{ji}^2 W_{jk}^2) + E(W_{ki}^2 W_{kj}^2)\}, \\
G_{IV} &= \sum_{1 \le i < j \le n} \{E(W_{ij} W_{ik} W_{lj} W_{lk}) + E(W_{ij} W_{il} W_{kj} W_{kl}) + E(W_{ik} W_{il} W_{jk} W_{jl})\}.
\end{aligned}
$$

Proof of theorem 2.1

    We can prove it by following the similar steps in the proof of Theorem 1 in Cai and Xu (2008) which derived the following Bahadur representation for any u_0,

$$\sqrt{nh} \begin{pmatrix} \beta_0 - a(u_0) \\ \beta_1 - a^{(1)}(u_0) \end{pmatrix} = \frac{(\Omega^*(u_0))^{-1}}{\sqrt{nh} f_u(u_0)} \sum_{t=1}^{n} \varphi_\tau(Y_i^*) X_i K(\frac{U_t - u_0}{h}) + o_p(\frac{1}{\sqrt{nh}}). \quad (2.18)$$

By using the leave-one-out estimation method, we can get the similar Bahadur representation for each design point U_t. All are similar except that X_t is not used

to estimate $A(U_t)$. So, we have

$$
\begin{aligned}
\hat{A}_{-t}(U_t) - A(U_t) &= \frac{(\Omega^*(U_t))^{-1}}{nh f_u(U_t)} \sum_{i \neq t}^{n} \varphi_\tau(Y_i^*) X_i K\left(\frac{U_i - U_t}{h}\right) + o_p\left(\frac{1}{\sqrt{nh}}\right) \\
&= R_1(U_t) + R_2(U_t) + o_p\left(\frac{1}{\sqrt{nh}}\right).
\end{aligned}
$$

and it holds uniformly for all U_t. $\square$

Proof of Theorem 2.2

The test statistic T_n under H_0 can be rewritten as

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau\left(Y_t - \hat{A}_{-t}(U_t)^T X_t\right) - \sum_{t=1}^{n} \rho_\tau\left(Y_t - A_0(U_t)^T X_t\right), \qquad (2.19)
\end{aligned}
$$

by applying Knight identity $\rho_\tau(r-s) - \rho_\tau(r) = s\{I(r < 0) - \tau\} + \int_0^s \{I(r < u) - I(r < 0)\} du$. Then, we can rewrite equation (2.19) as

$$
\begin{aligned}
T_n &= \sum_{t=1}^{n} ((\hat{A}_{-t}(U_t) - A_0(U_t))^T X_t (I_{(Y_t - A_0(U_t)^T X_t < 0)} - \tau) \\
&\quad + \sum_{t=1}^{n} \int_0^{((\hat{A}_{-t}(U_t) - A_0(U_t))^T X_t} (I_{(Y_t - A_0(U_t)^T X_t < s)} - I_{(Y_t - A_0(U_t)^T X_t < 0)}) ds \\
&= B_1 + B_2.
\end{aligned}
$$

First, we consider B_2.

$$
\begin{aligned}
B_2 &= \sum_{t=1}^{n} \int_0^{(\hat{A}_{-t}(U_t) - A_0(U_t))^T X_t} I_{(Y_t - A_0(U_t)^T X_t < s)} - I_{(Y_t - A_0(U_t)^T X_t < 0)} ds \\
&= E(B_2) + (B_2 - EB_2) \\
&= C_1 + C_2.
\end{aligned}
$$

For C_2, we can apply the law of large number to get $C_2 = o_p(1)$. For C_1, we define $\epsilon_t = Y_t - a(U_t)^T X_t$. By using the fact that ϵ_t is independent of $A_0(U_t)$ and $\hat{A}_{-t}(U_t)$, we can switch the order of integral and the expectation under condition $z = (u, x, y_{-t})$, where y_{-t} is $(Y_1, Y_2, ...Y_{t-1}, Y_{t+1}, ...Y_n)$,

$$
\begin{aligned}
C_1 &= \sum_{t=1}^{n} E\left(\int_0^{(\hat{A}_{-t}(U_t) - A_0(U_t))^T X_t} I_{Y_t - A_0(U_t)^T X_t < s} - I_{(Y_t - A_0(U_t)^T X_t < 0)} ds \right) \\
&= \sum_{t=1}^{n} E\left(\int_0^{(\hat{A}_{-t}(U_t) - A_0(U_t))^T X_t} E_z(I_{Y_t - A_0(U_t))^T X_t < s} - I_{(Y_t - A_0(U_t)^T X_t < 0)} | X_t, U_t) ds \right) \\
&= \sum_{t=1}^{n} E\left(\int_0^{(\hat{A}_{-t}(U_t) - A_0(U_t))^T X_t} E_z(I_{\epsilon_t < s} - I_{(\epsilon_t < 0)} | X_t, U_t) ds \right) \\
&= \sum_{t=1}^{n} E\left(\int_0^{(\hat{A}_{-t}(U_t) - A_0(U_t))^T X_t} F_{y|u,x}(s) - F_{y|u,x}(0) ds \right) \\
&= \frac{1}{2} \sum_{t=1}^{n} E(f_{y|u,x}(q_\tau | X_t, U_t))(\hat{A}_{-t}(U_t) - A_0(U_t))^T X_t X_t^T (\hat{A}_{-t}(U_t) - A_0(U_t)).
\end{aligned}
$$

we get

$$
\begin{aligned}
T_n &= -\sum_{t=1}^{n} (\hat{A}_{-t}(U_t) - A(U_t))^T X_t \varepsilon_t + \frac{1}{2} \sum_{t=1}^{n} f_{y|u,x}(q_\tau | X_t, U_t)(\hat{A}_{-t}(U_t) - A(U_t))^T \\
&\quad X_t X_t^T (\hat{A}_{-t}(U_t) - A(U_t)) + o_p(h^{-\frac{1}{2}}) \\
&= -\sum_{t=1}^{n} ((R_1(U_t) + R_2(U_t)^T X_t \varepsilon_t + \frac{1}{2} \sum_{t=1}^{n} f_{y|u,x}(q_\tau | X_t, U_t)(R_1(U_t) + R_2(U_t))^T \\
&\quad X_t X_t^T (R_1(U_t) + R_2(U_t)) + o_p(h^{-\frac{1}{2}})
\end{aligned}
$$

$$
\begin{aligned}
&= -\sum_{t=1}^{n} R_1(U_t)^T X_t \varepsilon_t - \sum_{t=1}^{n} R_2(U_t)^T X_t \varepsilon_t + \frac{1}{2} \sum_{t=1}^{n} f_{y|u,x}(q_\tau | X_t, U_t) R_1(U_t)^T \\
&\quad X_t X_t^T R_1(U_t) + \sum_{t=1}^{n} f_{y|u,x}(q_\tau | X_t, U_t) R_1(U_t)^T X_t X_t^T R_2(U_t) + \\
&\quad \frac{1}{2} \sum_{t=1}^{n} f_{y|u,x}(q_\tau | X_t, U_t) R_2^T(U_t) X_t X_t^T R_2(U_t) + o_p(h^{-\frac{1}{2}}) \\
&= -T_1 - T_2 + T_3 + T_4 + T_5 + o_p(h^{-\frac{1}{2}}).
\end{aligned}
$$

According to Cai and Xu (2008),we have the fact that:

$$R_2(U_t) = \frac{h^2}{2}a^{(2)}(U_t)\mu_2 + o_p(h^2).$$

Where $\mu_2 = \int t^2 K(t)dt$.

Therefore

$$
\begin{aligned}
T_2 &= \sum_{t=1}^{n} \varepsilon_t X_t \frac{h^2}{2}a^{(2)}(U_t)\mu_2 + o(nh^2), \\
T_4 &= \frac{h^2\mu_2}{n} \sum_{t=1}^{n} \frac{f_{y|u,x}(q_\tau(U_t,X_t))a^{(2)}(U_t)}{f_u(U_t)\Omega^*(U_t)} \sum_{j\neq t} \varepsilon_j X_j K_h(\frac{U_j - U_t}{h}) = O(h^2), \\
T_5 &= \frac{1}{8}\sum_{t=1}^{n} f_{y|u,x}(q_\tau(U_t,X_t)X_t X_t^T h^4 (a^{(2)}(U_t)\mu_2)^2 + o(nh^4).
\end{aligned}
$$

For T_1,we get

$$
\begin{aligned}
T_1 &= \frac{1}{nh}\sum_{t=1}^{n}(f_u(U_t)\Omega^*(U_t))^{-1} \sum_{i\neq t}\varepsilon_i X_i^T K(\frac{U_i - U_t}{h})X_t\varepsilon_t \\
&= \frac{1}{nh}\sum_{i\neq t}\varepsilon_t\varepsilon_i X_i^T(f_u(U_t)\Omega^*(U_t))^{-1}X_t K(\frac{U_i - U_t}{h}) \\
&= \frac{1}{n}\sum_{i\neq t}\varepsilon_t\varepsilon_i X_i^T(f_u(U_t)\Omega^*(U_t))^{-1}X_t K_h(U_i - U_t), \qquad (2.20)
\end{aligned}
$$

where $K_h(.)$ means $K(./h)/h$. Next we deal with T_3.

$$
\begin{aligned}
T_3 &= \frac{1}{2}\sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t,U_t))\frac{1}{nhf_u(U_t)}(\Omega^*(U_t))^{-1}\sum_{i\neq t}\varepsilon_i X_i^T K(\frac{U_i - U_t}{h})X_t X_t^T \\
&\quad \frac{1}{nhf_u(U_t)}(\Omega^*(U_t))^{-1}\sum_{j\neq t}\varepsilon_j X_j K(\frac{U_j - U_t}{h}) \\
&= \frac{1}{2n^2}\sum_{i=j}\varepsilon_i\varepsilon_j X_i^T\{\sum_{t\neq i}^{n} E(f_{y|u,x}(q_\tau|X_t,U_t))(f_u(U_t)\Omega^*(U_t))^{-1}X_t X_t^T \\
&\quad (f_u(U_t)\Omega^*(U_t))^{-1}K_h(U_i - U_t))K_h(U_j - U_t)X_j\}
\end{aligned}
$$

$$+ \frac{1}{2n^2} \sum_{i \neq j} \varepsilon_i \varepsilon_j X_i^T \{ \sum_{t \neq i,j}^n E(f_{y|u,x}(q_\tau | X_t, U_t))(f_u(U_t)\Omega^*(U_t))^{-1} X_t X_t^T$$

$$(f_u(U_t)\Omega^*(U_t))^{-1} K_h(U_i - U_t)) K_h((_j - U_t)\} X_j$$

$$= T_{31} + T_{32}.$$

by Lemma 1, we have

$$
\begin{aligned}
T_{31} &= \frac{1}{2} E(\varepsilon_i^2 X_i^T E(f_{y|u,x}(q_\tau | X_t, U_t))(f_u(U_t)\Omega^*(U_t))^{-1} X_t X_t^T \\
&\quad (f_u(U_t)\Omega^*(U_t))^{-1} X_i K_h^2(U_i - U_t)) \\
&= \frac{1}{2} tr\{ E(\varepsilon_i^2 E(f_{y|u,x}(q_\tau | X_t, U_t))(f_u(U_t)\Omega^*(U_t))^{-1} X_t X_t^T \\
&\quad (f_u(U_t)\Omega^*(U_t))^{-1} X_i X_i^T K_h^2(U_i - U_t)) \} \\
&= \frac{1}{2} \tau(1-\tau) tr\{ E[\frac{\Omega^*(U_i)^{-1}\Omega^*(U_t))}{f_u^2(U_t) f_{y|u,x}(q_\tau | X_i, U_i)} K_h^2(U_i - U_t)] \} \\
&= \frac{\tau(1-\tau)p}{2h} [E(\frac{1}{f_{y|u,x}(q_\tau | X_t, U_t)}) \int K^2(x) dx + O(h)].
\end{aligned}
$$

Next, we deal with the term T_{32},

$$
\begin{aligned}
T_{32} &= \frac{1}{2n} \sum_{i \neq j} \varepsilon_i \varepsilon_j X_i^T \frac{1}{n} \{ \sum_{t \neq i,j}^n E(f_{y|u,x}(q_\tau(U_t, X_t))(f_u(U_t)\Omega^*(U_t))^{-1} X_t X_t^T \\
&\quad (f_u(U_t)\Omega^*(U_t))^{-1} K_h(U_i - U_t) K_h(U_j - U_t) \} X_j \\
&= \frac{1}{2n} \sum_{i \neq j} Z(i,j) \varepsilon_i \varepsilon_j,
\end{aligned}
$$

Where

$$
\begin{aligned}
Z(i,j) &= \frac{1}{n} X_i^T \{ \sum_{t \neq i,j}^n E(f_{y|u,x}(q_\tau(U_t, X_t)))(f_u(U_t)\Omega^*(U_t))^{-1} X_t X_t^T (f_u(U_t)\Omega^*(U_t))^{-1} \\
&\quad K_h(U_i - U_t) K_h(U_j - U_t) \} X_j.
\end{aligned}
$$

For $i \neq j$, define

$$Q(i,j) = \sum_{i \neq j} \int_{\Omega} K_h(U_i - u) K_h(U_j - u) du X_i^T (f_u(U_j) \Omega^*(U_j))^{-1} X_j,$$

Let

$$Zn = \sum_{i \neq j} Z(i,j) \varepsilon_i \varepsilon_j,$$

$$Qn = \sum_{i \neq j} Q(i,j) \varepsilon_i \varepsilon_j,$$

Consider

$$E(Z_n - Q_n)^2 = E(Z_n^2) + E(Q_n^2) - 2E(Z_n Q_n),$$

For $E(Z_n^2)$, since $Z(i,j)$ is symmetric in (i,j), by lemma 1, we get

$$
\begin{aligned}
E(Z_n^2) &= 2(\tau(1-\tau))^2 \sum_{i \neq j} E(Z(i,j))^2 \\
&= \frac{2(\tau(1-\tau))^2}{n^2} \sum_{i \neq j} E[\sum_{t \neq i \neq j} (E(f_{y|u,x}(q_\tau(U_t, X_t))))^2 (X_i^T(f_u(U_t)\Omega^*(U_t))^{-1}X_t)^2 \\
&\quad (X_t^T(f_u(U_t)\Omega^*(U_t))^{-1}X_j)^2 K_h^2(U_i - U_t) K_h^2(U_t - U_j) \\
&\quad + \sum_{t,l \neq i \neq j} E_z(f_{y|u,x}(q_\tau(U_t, X_t))) E(f_{y|u,x}(q_\tau(U_l, X_l))) X_i^T(f_u(U_t)\Omega^*(U_t))^{-1} \\
&\quad (X_t X_t^T(f_u(U_t)\Omega^*(U_t))^{-1}X_j X_i^T(f_u(U_l)\Omega^*(U_l))^{-1}X_l X_l^T(f_u(U_l)\Omega^*(U_l))^{-1}X_j \\
&\quad K_h(U_i - U_t) K_h(U_t - U_j) K_h(U_i - U_l) K_h(U_l - U_j)] \\
&= \frac{2(\tau(1-\tau))^2}{n^2}[n^3 J_1 + n^4 J_2](1 + O(n^{-1}),
\end{aligned}
$$

where

$$
\begin{aligned}
J_1 &= E\{tr[E((f_{y|u,x}(q_\tau(U_t, X_t))))^2 X_i X_i^T(f_u(U_t)\Omega^*(U_t))^{-1} X_t X_t^T(f_u(U_t)\Omega^*(U_t))^{-1} \\
&\quad X_j X_j^T(f_u(U_t)\Omega^*(U_t))^{-1} X_t X_t^T(f_u(U_t)\Omega^*(U_t))^{-1} K_h^2(U_i - U_t) K_h^2(U_t - U_j) \\
&= tr\{E[\frac{\Omega^*(U_i)\Omega^*(U_j)}{f_u^4(U_t) f_{y|u,x}(q_\tau(U_i, X_i))) f_{y|u,x}(q_\tau(U_j, X_j)))\Omega^*(U_t))\Omega^*(U_t))}
\end{aligned}
$$

$$K_h^2(U_i - U_t)K_h^2(U_t - U_j)]\}$$

$$= \int_\Omega \int_\Omega \int_\Omega \frac{\Omega^*(U_i)\Omega^*(U_j)}{f_u^4(U_t)f_{y|u,x}(q_\tau(U_i, X_i)))f_{y|u,x}(q_\tau(U_j, X_j)))\Omega^*(U_t))\Omega^*(U_t))}$$
$$K_h^2(U_i - U_t)K_h^2(U_t - U_j)f_u(U_t)f_u(U_i)f_u(U_j)dU_t dU_i dU_j$$

$$= \frac{p}{h^2}[\int_\Omega \frac{1}{f_u(U_t)f_{y|u,x}^2(q_\tau(U_t, X_t))}dU_t \int K^2(x)dx \int K^2(y)dy + O(h)],$$

and

$$J_2 = E\{tr[E(f_{y|u,x}(q_\tau(U_t, X_t)))E_z(f_{y|u,x}(q_\tau(U_l, X_l)))X_i^T(f_u(U_t)\Omega^*(U_t))^{-1}$$

$$X_t X_t^T(f_u(U_t)\Omega^*(U_t))^{-1}X_j X_i^T(f_u(U_l)\Omega^*(U_l))^{-1}X_l X_l^T(f_u(U_l)\Omega^*(U_l))^{-1}X_j]$$

$$K_h(U_i - U_t)K_h(U_t - U_j)K_h(U_i - U_l)K_h(U_l - U_j)\}$$

$$= tr\{E[\frac{\Omega^*(U_i)\Omega^*(U_j)}{f_u^4(U_t)f_{y|u,x}(q_\tau(U_i, X_i)))f_{y|u,x}(q_\tau(U_j, X_j)))\Omega^*(U_l))\Omega^*(U_t))}$$
$$K_h(U_i - U_t)K_h(U_t - U_j)K_h(U_i - U_l)K_h(U_l - U_j)\}$$

$$= \int_\Omega \int_\Omega \int_\Omega \frac{\Omega^*(U_i)^{-1}\Omega^*(U_j)^{-1}}{f_u^4(U_t)f_{y|u,x}(q_\tau(U_i, X_i)))f_{y|u,x}(q_\tau(U_j, X_j)))\Omega^*(U_t))^{-1}\Omega^*(U_t))^{-1}}$$

$$K_h(U_i - U_t)K_h(U_t - U_j)K_h(U_i - U_l)K_h(U_l - U_j)\}$$

$$f_u(U_t)f_u(U_l)f_u(U_i)f_u(U_j)dU_t dU_i dU_j dU_l$$

$$= \frac{p}{h}[\int_\Omega \frac{1}{f_{y|u,x}^2(q_\tau(U_t, X_t))}dU_t \int \int \int K(x)K(y)K(z - x)K(z - y)dzdxdy$$
$$+ O(h)].$$

Next for $E(Q_n^2)$,

$$E(Q_n^2) = (\tau(1 - \tau))^2 \sum_{i \neq j} E(Q(i, j)^2) + (\tau(1 - \tau))^2 \sum_{i \neq j} E(Q(i, j)Q(j, i))$$

$$= n(n - 1)(\tau(1 - \tau))^2 E\{tr[X_i X_i^T(f_u(U_j)\Omega^*(U_j))^{-1}X_j X_j^T$$

$$(f_u(U_j)\Omega^*(U_j))^{-1}][\int_\Omega K_h(U_i - u)K_h(U_j - u)du]^2\} + n(n - 1)$$

$$(\tau(1 - \tau))^2 E\{tr[X_i X_i^T(f_u(U_i)\Omega^*(U_i))^{-1}X_j X_j^T(f_u(U_j)\Omega^*(U_j))^{-1}]$$

$$[\int_\Omega K_h(U_i-u)K_h(U_j-u)du]^2\}$$

$$
\begin{aligned}
= \quad & (\tau(1-\tau))^2 tr[E\frac{\Omega^*(U_i)}{f_u^2(U_j)(f_{y|u,x}(q_\tau(U_i,X_i)))\Omega^*(U_i))f_{y|u,x}(q_\tau(U_j,X_j))\Omega^*(U_j)} \\
& [\int_\Omega K_h(U_i-u)K_h(U_j-u)du]^2]n(n-1) \\
& + n(n-1)(\tau(1-\tau))^2 tr[E\frac{1}{f_u^2(U_j)(f_{y|u,x}(q_\tau(U_i,X_i)))f_{y|u,x}(q_\tau(U_j,X_j)))\Omega^*(U_j)} \\
& [\int_\Omega K_h(U_i-u)K_h(U_j-u)du]^2] \\
= \quad & \frac{2n^2(\tau(1-\tau))^2 p}{h}[E(\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))})+O(h)],
\end{aligned}
$$

$$
\begin{aligned}
E(Z_nQ_n) \quad = \quad & (\tau(1-\tau))^2\sum_{i\neq j}E(Z(i,j)(Q(i,j)+Q(j,i)] \\
= \quad & \frac{(\tau(1-\tau))^2}{n}E\{tr[\sum_{t\neq i\neq j}E_z(f_{y|u,x}(q_\tau(U_t,X_t)))X_tX_t^T(f_u(U_t)\Omega^*(U_t))^{-1} \\
& X_jX_j^T((f_u(U_j)\Omega^*(U_j))^{-1}+(f_u(U_i)\Omega^*(U_i))^{-1})X_iX_i^T \\
& (f_u(U_t)\Omega^*(U_t))^{-1}]K_h(U_i-U_t)K_h(U_t-U_j) \\
& \int_\Omega K_h(U_i-u)K_h(U_j-u)\} \\
= \quad & \frac{(\tau(1-\tau))^2}{n}\sum_{i\neq j}tr\{E[\frac{\Omega^*(U_i)\Omega^*(U_j)}{f_u^2(U_t)f_{y|u,x}(q_\tau(U_i,X_i)))f_{y|u,x}(q_\tau(U_j,X_j)))\Omega^*(U_t))} \\
& ((f_u(U_j)\Omega^*(U_j))^{-1} \\
& + (f_u(U_i)\Omega^*(U_i))^{-1})K_h(U_i-U_t)K_h(U_t-U_j)\int_\Omega K_h(U_i-u)K_h(U_j-u)\} \\
= \quad & \frac{2n^2(\tau(1-\tau))^2 p^2}{h}[E(\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))})+O(h)].
\end{aligned}
$$

So we have

$$E(Z_n - Q_n)^2 = E(Z_n^2) + E(Q_n)^2 - 2E(Z_nQ_n) = O(n^2) + O(nh^{-2}).$$

Then as $h \to 0, nh \to \infty$, we have

$$\sum_{i \neq j} Z(i,j)\varepsilon_i\varepsilon_j = \sum_{i \neq j} \int_\Omega K_h(U_i - U_t)K_h(U_j - U_t)duX_i^T(f_u(U_t)\Omega^*(U_t))^{-1}X_j\varepsilon_i\varepsilon_j$$
$$+ O(n) + O(\sqrt{n}h^{-1}),$$

$$T_{32} = \frac{1}{2n}\sum_{i \neq j}\varepsilon_i\varepsilon_j \int_\Omega K_h(U_i - U_t)K_h(U_j - U_t)duX_i^T(f_u(U_t)\Omega^*(U_t))^{-1}X_j + O(1)$$
$$+ O(\frac{1}{\sqrt{nh}}),$$

So,

$$T_3 = \frac{1}{2h}p\tau(1-\tau)E(\frac{1}{f_{y|u,x}(q_\tau(U_t, X_t))})\int K^2(t)dt + o_p(h^{-1/2})$$
$$+ \frac{1}{2n}\sum_{i<j}\varepsilon_i\varepsilon_j X_i^T(f_u(U_t)\Omega^*(U_t))^{-1}K_h * K_h((U_i - U_j)/h)X_j + O(1)$$
$$+ O(\frac{1}{\sqrt{nh}}),$$

$$-T_1 + T_3 = -\frac{1}{n}\sum_{i \neq t}\varepsilon_t\varepsilon_i X_i^T(f_u(U_t)\Omega^*(U_t))^{-1}X_tK_h(\frac{U_i - U_t}{h})$$
$$+ \frac{1}{2h}p\tau(1-\tau)E(\frac{1}{f_{y|u,x}(q_\tau(U_t, X_t))})\int K^2(t)dt + o_p(h^{-1/2})$$
$$+ \frac{1}{2n}\sum_{i<j}\varepsilon_i\varepsilon_j X_i^T(f_u(U_t)\Omega^*(U_t))^{-1}K_h * K_h((U_i - U_j)/h)X_j$$
$$+ O(1) + O(\frac{1}{\sqrt{nh}})$$
$$= \frac{1}{2h}p\tau(1-\tau)E(\frac{1}{f_{y|u,x}(q_\tau(U_t, X_t))})\int K^2(t)dt - W(n)h^{-\frac{1}{2}}/2$$

$$+ O(1) + O(\frac{1}{\sqrt{nh}}),$$

where

$$W(n) = \frac{\sqrt{h}}{n} \sum_{j \neq l} \varepsilon_i \varepsilon_l [2K_h(U_j - U_l) - K_h * K(U_j - U_l)] X_j^T(q_\tau)(f_u(U_t)\Omega^*(U_t))^{-1} X_l.$$

let $v = 2||2K - K * K||_2^2 \frac{(\tau(1-\tau))^2 p}{h} E(\frac{1}{f_{y|u,x}^2(q_\tau(U_t, X_t))})$, I will show $var(W(n)) \to v$, and then I can apply lemma 3 to get $W(n) \to N(0, v)$.

To prove it, we define $W_{i,j} = \frac{\sqrt{h}}{n} w(i, j)\varepsilon_i \varepsilon_j$, where $w(i, j)$ is written in a symmetric form:

$$w(i, j) = w_1(i, j) + w_2(i, j) - w_3(i, j) - w_3(i, j),$$

where

$$
\begin{aligned}
w_1(i, j) &= 2K_h(U_i - U_j)X_i^T f_u(U_j)\Omega^*(U_j))^{-1} X_j, \\
w_2(i, j) &= w_1(j, i), \\
w_3(i, j) &= \int_\Omega K_h(U_i - u)K_h(U_j - u)X_i^T(f_u(U_j)\Omega^*(U_j))^{-1} X_j, \\
w_4(i, j) &= w_3(j, i).
\end{aligned}
$$

then $W(n) = \sum_{i<j} W(i, j)$ and $var(W(n)) = \sum_{i<j} E(W(i, j)^2)$.

To apply lemma 3, we need to verify the following conditions.

Condition 1) $W(n)$ is clean

Condition 2) $var(W(n)) \to v$.

Condition 3) G_I is of smaller order than $var(W(n))$.

Condition 4) G_{II} is of smaller order than $var(W(n))$.

Condition 5) G_{IV} is of smaller order than $var(W(n))$.

where G_I, G_{II} and G_{IV} are defined in the lemma 3.

I will show the proof in detail as follows

Condition 1)   by the definition of clean and $W(n)$, the proof is obvious.

Condition 2)   We note that

$$var(W(n)) = \sum_{i<j} EW_{ij}^2,$$

By the straightforward calculation,

$$
\begin{aligned}
E(w_1^2(i,j)\varepsilon_i^2\varepsilon_j^2) &= \frac{4(\tau(1-\tau))^2 p}{h} E\big[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}\big] \int_\Omega K^2(y)dy \\
&\quad (1+O(h)), \\
E(w_2^2(i,j)\varepsilon_i^2\varepsilon_j^2) &= \frac{4(\tau(1-\tau))^2 p}{h} E\big[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}\big] \int_\Omega K^2(y)dy \\
&\quad (1+O(h)), \\
E(w_3^2(i,j)\varepsilon_i^2\varepsilon_j^2) &= \frac{(\tau(1-\tau))^2 p}{h} E\big[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}\big] \\
&\quad (\int_\Omega \int_\Omega K(y-x)K(x)dxdy)^2(1+O(h)), \\
E(w_4^2(i,j)\varepsilon_i^2\varepsilon_j^2) &= \frac{(\tau(1-\tau))^2 p}{h} E\big[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}\big] \\
&\quad (\int_\Omega \int_\Omega K(y-x)K(x)dxdy)^2(1+O(h)), \\
E(w_1(i,j)w_2(i,j)\varepsilon_i^2\varepsilon_j^2) &= \frac{4(\tau(1-\tau))^2 p}{h} E\big[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}\big] \\
&\quad \int_\Omega K^2(y)dy(1+O(h)), \\
E(w_1(i,j)w_3(i,j)\varepsilon_i^2\varepsilon_j^2) &= \frac{2(\tau(1-\tau))^2 p}{h} E\big[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}\big] \\
&\quad \int_\Omega K(y)dy \int_\Omega K(u)K(x+u)dx(1+O(h)), \\
E(w_1(i,j)w_4(i,j)\varepsilon_i^2\varepsilon_j^2) &= \frac{2(\tau(1-\tau))^2 p}{h} E\big[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}\big] \\
&\quad \int_\Omega K(y)dy \int_\Omega K(u)K(x+u)dx(1+O(h)),
\end{aligned}
$$

$$E(w_2(i,j)w_3(i,j)\varepsilon_i^2\varepsilon_j^2) = \frac{2(\tau(1-\tau))^2 p}{h} E[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}]$$

$$\int_\Omega K(y)dy \int_\Omega K(u)K(x+u)dx(1+O(h)),$$

$$E(w_2(i,j)w_2(i,j)\varepsilon_i^2\varepsilon_j^2) = \frac{2(\tau(1-\tau))^2 p}{h} E[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}]$$

$$\int_\Omega K(y)dy \int_\Omega K(u)K(x+u)dx(1+O(h)),$$

$$E(w_3(i,j)w_4(i,j)\varepsilon_i^2\varepsilon_j^2) = \frac{(\tau(1-\tau))^2 p}{h} E[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}]$$

$$(\int_\Omega \int_\Omega K(y-x)K(x)dxdy)^2(1+O(h)).$$

So,

$$E(w(i,j)^2\varepsilon_i^2\varepsilon_j^2) = \frac{(2\tau(1-\tau))^2 p}{h} \int [2K(x)-K*K(x)]^2 dx E[\frac{1}{f_{y|u,x}^2(q_\tau(U_j,X_j))}].$$

Therefore, $v = 2\|2K - K*K\|_2^2 \frac{(\tau(1-\tau))^2 p^2}{h} E(\frac{1}{f_{y|u,x}^2(q_\tau(U_t,X_t))})$.

Condition 3) Note that $E(w_1(i,j)\varepsilon_i\varepsilon_j)^4 = O(h^{-3})$ and $E(w_3(i,j)\varepsilon_i\varepsilon_j)^4 = O(h^{-2})$, therefore $E(w(i,j)\varepsilon_i\varepsilon_j)^4 = \frac{h^2}{n^4}O(h^{-3})$. Hence we get $G_I = O(n^{-2}h^{-1}) = o(1)$.

Condition 4) $E(w(i,j)^2 w(j,k)^2) = O(n^{-4}h^{-1})$ which lead to $G_{II} = O(n^3 \frac{1}{n^4 h}) = O(\frac{1}{nh}) = o(1)$.

Condition 5) By the definition of symmetric form of $w_i i, j$, it suffices to consider this term $E(W_{ij}W_{ik}W_{lj}W_{lk})$. Note that

$$E\{w_1(i,j)w_1(i,k)w_1(l,j)w_1(l,k)\} = O(h^{-1}),$$

$$E\{w_1(i,j)w_1(i,k)w_1(l,j)w_3(l,k)\} = O(h^{-1}),$$

$$E\{w_1(i,j)w_1(i,k)w_3(l,j)w_3(l,k)\} = O(h^{-1}),$$

$$E\{w_1(i,j)w_3(i,k)w_3(l,j)w_3(l,k)\} = O(h^{-1}),$$

$$E\{w_3(i,j)w_3(i,k)w_3(l,j)w_3(l,k)\} = O(h^{-1}).$$

Therefore, we get $E(W_{ij}W_{ik}W_{lj}W_{lk}) = O(\frac{h^2}{n^4})O(h^{-1}) = O(n^{-4}h)$. Then we get

$$G_{IV} = O(h) = o(1).$$

Then, all conditions for Lemma 3 are satisfied. By applying Lemma 3 we can get the asymptotic distribution which finishes the proof of Theorem 2.2.□

Proof of Theorem 2.3

Under $H_0 : A(u) = A_0$, we have $A_0^{(1)}(U_t) = 0$ since it is a constant vector. So $Y_t^* = Y_t - X_t'A_0(U_t)$, which implies that $\varphi(Y_t^*) = \varepsilon_t$. Therefore, we obtain $R_2(U_t) = 0$. Thus,

$$
\begin{aligned}
\hat{A}_{-t}(U_t) - A_0(U_t) &= \frac{(\Omega^*(U_t))^{-1}}{nhf_u(U_t)} \sum_{i \neq t}^{n} \varphi_\tau(Y_i^*)X_i K(\frac{U_i - U_t}{h}) + o_p(\frac{1}{\sqrt{nh}}) \\
&= R_1(U_t) + o_p(\frac{1}{\sqrt{nh}}).
\end{aligned}
$$

Under H_0, the test statistic T_n can be rewritten as

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_{-t}(U_t)^T X_t\right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_0^T(U_t)X_t\right),
\end{aligned}
$$

where A_0 is the constant vector. Denote $\varepsilon_t = \tau - I_{(Y_t - A_0(U_t)^T X_t < 0)}$, and recall $R_2(U_t) = 0$ under H_0. By applying Knight identity and the similar derivation in the proof of Theorem 2.2, then, we get

$$
\begin{aligned}
T_n &= -\sum_{t=1}^{n}(\hat{A}_{-t}(U_t) - A_0(U_t))^T X_t \varepsilon_t \\
&\quad + \frac{1}{2}\sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t))(\hat{A}_{-t}(U_t) - A_0(U_t))^T X_t X_t^T(\hat{A}_{-t}(U_t) - A_0(U_t))
\end{aligned}
$$

$$
\begin{aligned}
= \quad & -\sum_{t=1}^{n}((R_1(U_t) + R_2(U_t)^T X_t \varepsilon_t \\
& + \frac{1}{2}\sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t))(R_1(U_t) + R_2(U_t))^T X_t X_t^T (R_1(U_t) + R_2(U_t)) \\
= \quad & -\sum_{t=1}^{n} R_1(U_t)^T X_t \varepsilon_t + \frac{1}{2}\sum_{t=1}^{n} E_z(f_{y|u,x}(q_\tau|X_t, U_t)) R_1(U_t)^T X_t X_t^T R_1(U_t) \\
= \quad & -T_1 + T_2.
\end{aligned}
$$

Next, following the similar derivation in the proof of Theorem 2.2, we can prove $W(n) \to N(0, v)$ with $v = 2\|2K - K * K\|_2^2 p E(\frac{1}{f_{y|u,x}(q_\tau(U_t, X_t))f_u(U_t)}).\square$

Proof of Theorem 2.4

Under H_0, we have $A^{(1)}(U_t) = 0$, So $Y_t^* = Y_t - X_t' A(U_t)$, $R_2(U_t) = 0$, and

$$
\begin{aligned}
\hat{A}_{-t}(U_t) - A(U_t) &= \frac{(\Omega^*(U_t))^{-1}}{(\sqrt{nh})^2 f_u(U_t)} \sum_{i \neq t}^{n} \varphi_\tau(Y_i^*) X_i K(\frac{U_i - U_t}{h}) + o_p(\frac{1}{\sqrt{nh}}) \\
&= R_1(U_t) + o_p(\frac{1}{\sqrt{nh}}).
\end{aligned}
$$

Then, the test statistic T_n can be rewritten as

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau\left(Y_t - \hat{A}_{-t}(U_t)^T X_t\right) - \sum_{t=1}^{n} \rho_\tau\left(Y_t - \hat{A}_0^T X_t\right) \\
&= \sum_{t=1}^{n} \rho_\tau\left(Y_t - \hat{A}_{-t}(U_t)^T X_t\right) - \sum_{t=1}^{n} \rho_\tau\left(Y_t - A_0(U_t)^T X_t\right) \\
&\quad + \sum_{t=1}^{n} \rho_\tau\left(Y_t - A_0(U_t)^T X_t\right) - \sum_{t=1}^{n} \rho_\tau\left(Y_t - \hat{A}_0^T X_t\right) \\
&= B_1 + B_2,
\end{aligned}
$$

where

$$B_1 = \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_{-t}(U_t)^T X_t \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_0(U_t)^T X_t \right),$$

$$B_2 = \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_0(U_t)^T X_t \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_0^T X_t \right).$$

For B_2, according to Koenker (2005), we know the limiting distribution of B_2 would be a χ^2 distribution under H_0. Therefore, $B_2 = O_p(1)$. Then, the asymptotic distribution of the B_1 can be derived by using the similar procedure as that for proving Theorem 2.2.

Proof of Theorem 2.6

From theorem 2.5, we have known $|\hat{f}_{y|u,x}(q_\tau(U_t, X_t)) - f_{y|u,x}(q_\tau(U_t, X_t))| = o_p(1)$, so it is obvious that

$$|\hat{\mu}_n - \mu_n| = o_p(1),$$
$$|\hat{\sigma}_n^2 - \sigma_n^2| = o_p(1),$$
$$|\hat{d}_n - d_n| = o_p(1).$$

Then we are done by following the proof of theorem 2.3. $\square$

Proof of Theorem 2.7

Under the H_a,the true parameter $A(u) = A_0(U) + \frac{1}{\sqrt{nh}}\Delta(U)$. The test statistic T_n can be rewritten as

$$T_n = \ell(H_a) - \ell(H_0)$$
$$= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_{-t}(U_t)^T X_t \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}^T X_t \right)$$

$$\begin{aligned}
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_{-t}(U_t)^T X_t\right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - A(U_t)^T X_t\right) \\
&\quad + \sum_{t=1}^{n} \rho_\tau \left(Y_t - A(U_t)^T X_t\right)) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_0^T X_t\right) \\
&= E_1 + E_2.
\end{aligned}$$

First, I consider E_2, by using Knight identity

$$\begin{aligned}
E_2 &= \sum_{t=1}^{n} \rho_\tau \left(Y_t - A(U_t)^T X_t\right)) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_0^T X_t\right) \\
&= -(\sum_{t=1}^{n} \rho_\tau \left(Y_t - A_0^T X_t\right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - A(U_t)^T X_t\right))) \\
&= -\frac{1}{\sqrt{nh}} \sum_{t=1}^{n} \Delta(U_t) X_t \varepsilon_t - \int_{0}^{-\frac{1}{\sqrt{nh}}\Delta(U_t)X_t} I_{\epsilon_t < s} - I_{\epsilon_t < 0} ds \\
&= -F_1 - F_2.
\end{aligned}$$

By following the similar derivation in theorem 2.2, we can get

$$\begin{aligned}
F_2 &= \frac{1}{2nh} \sum_{t=1}^{n} E_z(f_{y|u,x}(q_\tau|X_t, U_t)) \Delta(U_t)^T X_t X_t^T \Delta(U_t) \\
&= \frac{1}{2h} E(f_{y|u,x}(q_\tau|X_t, U_t) \Delta(U_t)^T X_t X_t^T \Delta(U_t)) \\
&\equiv d_{2n}.
\end{aligned}$$

For E_1 we have the same result as theorem 2.2 except that

$$R_2(U_t) = \frac{h^2}{2} (\frac{1}{\sqrt{nh}} \Delta^{(2)}(u)) \mu_2 + o_p(h^2).$$

Therefore,

$$E_1 = -\sum_{t=1}^{n} (\hat{A}_{-t}(U_t) - A(U_t))^T X_t \varepsilon_t$$

$$+ \frac{1}{2} \sum_{t=1}^{n} E_z(f_{y|u,x}(q_\tau|X_t, U_t))(\hat{A}_{-t}(U_t) - A(U_t))^T X_t X_t^T (\hat{A}_{-t}(U_t) - A(U_t))$$

$$= -\sum_{t=1}^{n} ((R_1(U_t) + R_2(U_t)^T X_t \varepsilon_t$$

$$+ \frac{1}{2} \sum_{t=1}^{n} E_z(f_{y|u,x}(q_\tau|X_t, U_t))(R_1(U_t) + R_2(U_t))^T X_t X_t^T (R_1(U_t) + R_2(U_t))$$

$$= -\sum_{t=1}^{n} R_1(U_t)^T X_t \varepsilon_t - \sum_{t=1}^{n} R_2(U_t)^T X_t \varepsilon_t$$

$$+ \frac{1}{2} \sum_{t=1}^{n} E_z(f_{y|u,x}(q_\tau|X_t, U_t)) R_1(U_t)^T X_t X_t^T R_1(U_t)$$

$$+ \sum_{t=1}^{n} E_z(f_{y|u,x}(q_\tau|X_t, U_t)) R_1(U_t)^T X_t X_t^T R_2(U_t)$$

$$+ \frac{1}{2} \sum_{t=1}^{n} f_{y|u,x}(q_\tau|X_t, U_t) R_2^T(U_t) X_t X_t^T R_2(U_t)$$

$$= -O_1 - O_2 + O_3 + O_4 + O_5.$$

So, under H_a we have

$$O_2 = \frac{1}{\sqrt{nh}} \sum_{t=1}^{n} \varepsilon_t X_t \frac{h^2}{2} \Delta^{(2)}(U_t) \mu_2 + o(nh^2),$$

$$O_4 = \frac{1}{\sqrt{nh}} \frac{h^2 \mu_2}{n} \sum_{t=1}^{n} \frac{f_{y|u,x}(q_\tau(U_t, X_t)) \Delta^{(2)}(U_t)}{f_u(U_t) \Omega^*(U_t)} \sum_{j \neq t} \varepsilon_j X_j K_h(\frac{U_j - U_t}{h}),$$

$$O_5 = \frac{1}{8nh} \sum_{t=1}^{n} f_{y|u,x}(q_\tau(U_t, X_t) X_t X_t^T h^4 (\Delta^{(2)}(U_t) \mu_2)^2 + o(nh^4).$$

So, $O_2 + O_4 + O_5$ is dominated by d_{2n}, which is the $O(1/h)$.

For O_1 and O_3 we have

$$-O_1 + O_3 = -\frac{1}{n} \sum_{i \neq t} \varepsilon_t \varepsilon_i X_i^T (f_u(U_t) \Omega^*(U_t))^{-1} X_t K_h(\frac{U_i - U_t}{h})$$

$$+ \frac{1}{2h} p\tau(1-\tau) E(\frac{1}{f_{y|u,x}(q_\tau(U_t, X_t))}) \int K^2(t) dt + o_p(h^{-1/2})$$

$$+ \frac{1}{nh} \sum_{i<j} \varepsilon_i \varepsilon_j X_i^T (f_u(U_t) \Omega^*(U_t))^{-1} K * K((U_i - U_j)/h) X_j$$

$$+ O(1) + O(\frac{1}{\sqrt{nh}})$$

$$= \frac{1}{2h} p\tau(1-\tau) E(\frac{1}{f_{y|u,x}(q_\tau(U_t, X_t))}) \int K^2(t)dt - W(n)h^{-\frac{1}{2}}/2$$

$$+ O(1) + O(\frac{1}{\sqrt{nh}}).$$

$$T_n = \mu_n - O_2 + O_4 + O_5 - (W(n)h^{-\frac{1}{2}}/2 + \frac{1}{\sqrt{nh}} \sum_{t=1}^{n} \Delta(U_t) X_t \varepsilon_t) + O(1) + O(\frac{1}{\sqrt{nh}}).$$

where $W(n)$ is defined in the proof of theorem 2.2. The rest of proof is similar to the proof of theorem 2.2. The details are omitted. $\square$

CHAPTER 3: GENERALIZED QUASI-LIKELIHOOD TEST OF PARTIAL
COEFFICIENTS FOR VARYING COEFFICIENT QUANTILE REGRESSION
MODEL

The focus for this chapter is mainly on testing whether partial coefficients in varying coefficient quantile regression models are constant or of some specific functional forms with other coefficients completely unspecified.

3.1 Introduction

Chapter 2 is devoted to proposing and studying the generalized quasi-likelihood ratio test statistic to check whether all coefficients in varying coefficient quantile regression model are constant or of particular form. In this chapter, I will apply the proposed test procedure to check whether partial coefficients in varying coefficient quantile regression model are constant or of some specific functional form with other coefficients remaining completely unspecified. In other words, under the null hypothesis, the model becomes a partially varying coefficient quantile model. These hypotheses are also of considerable interest due to various applications. By performing these tests, an accurate form of the model for a real data analysis can be obtained. For example, if we know partial coefficients are constant for a real application, we just use a partially varying coefficient quantile regression model to allow for appreciable flexibility on the structure of the fitted model since it admits some coefficients to be constant and the others to be functional. A partially varying coefficient quantile regression model serves as an intermediate class between the fully nonparametric models as in Honda (2004) and Cai and Xu (2008) and the fully parametric models as in Koenker and Xiao (2006). Besides good properties such as

robustness by nonparametric treatment on certain variables, it also can provide an efficient estimation on the parametric effect of other variables.

Recall that a varying coefficient quantile regression model takes the following form

$$q_\tau(U_t, X_t) = \sum_{k=0}^{p} a_{k,\tau} X_{tk} = A(U_t)^T X_t \equiv A_1(U_t)^T X_t^{(1)} + A_2(U_t)^T X_t^{(2)}, \qquad (3.1)$$

where U_t is the smoothing variable, $A(\cdot) = (A_1(\cdot)^T, A_2(\cdot)^T)^T$ and $X_t = (X_{t1}^T, X_{t2}^T)^T$ with $A_1(\cdot)$ and $X^{(1)}$ being $p_1(< p)$ dimensional. In this case, our interest is to test whether $A_1(\cdot)$ is indeed a constant vector or of some specific functional form with $A_2(\cdot)$ remaining completely unspecified. This is equivalent to performing the following testing hypothesis

$$H_0 : A_1(u) = A_{10} \quad \text{versus} \quad H_a : A_1(u) \neq A_{10} \qquad (3.2)$$

with A_{10} is a known or unknown constant vector with $A_2(u)$ remaining completely unspecified, and

$$H_0 : A_1(u) = A_{10}(u) \quad \text{versus} \quad H_a : A_1(u) \neq A_{10}(u), \qquad (3.3)$$

where $A_{10}(u)$ is a vector of known functionals with $A_2(u)$ remaining completely unspecified.

The rest of this chapter is organized as follows. I present the generalized likelihood ratio test statistics to the above hypotheses in (3.2) and (3.3), respectively and derive their asymptotic properties in the following subsections. Also, I investigate the power of the proposed test statistics. Finally, I will conduct Monte Carlo simulations to illustrate the finite sample performance and a real application is considered to demonstrate the effectiveness of the proposed methods.

3.2 Test Statistic and Its Asymptotic Distribution

3.2.1 Test of Functional Form of Partial Varying Coefficient

First, we consider the hypothesis given in (3.3). Following the same derivations in Chapter 2, I can construct the quasi-likelihood function under both the null and alternative hypotheses by estimating the unknown coefficients using local linear technique with jackknife method. Note that it does not need to estimate $A_1(\cdot)$ under the null hypothesis since $A_{10}(\cdot)$ is known. Let $Y_t^* = Y_t - A_{10}(U_t)X_t^{(1)}$ as the partial residual in the estimation of $A_2(\cdot)$ under the null hypothesis. The generalized quasi-likelihood ratio (GQLR) test statistic can be defined as

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^n \rho_\tau \left(Y_t - \hat{A}_1(U_t)X_t^{(1)} - \hat{A}_2(U_t)X_t^{(2)} \right) - \sum_{t=1}^n \rho_\tau \left(Y_t^* - \hat{A}_2^*(U_t)X_t^{(2)} \right),
\end{aligned} \tag{3.4}
$$

where $\hat{A}_1(\cdot)$ and $\hat{A}_2(\cdot)$ are the nonparametric estimators from model (3.1) by using the estimation procedure outlined in Chapter 2; see (2.6) for details, and $\hat{A}_2^*(\cdot)$ is the nonparametric estimator from the varying coefficient quantile model as $q_\tau(Y^*|U, X^{(2)}) = A_2^*(U)^T X^{(2)}$ by using the same estimation procedure as earlier; again, see (2.6) for details.

Recall that $\Omega^*(U_t) = E(X_t X_t^T f_{y|u,x}(q_\tau(U_t, X_t))|U_t)$. Rewrite the matrix as follows

$$
\Omega^*(U_t) = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}, \quad \text{and} \quad \Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21},
$$

where Ω_{11}, Ω_{12}, Ω_{21}, and Ω_{22} are $p_1 \times p_1$, $p_1 \times p_2$, $p_2 \times p_1$, $p_2 \times p_2$ matrices and $p_2 = p - p_1$, respectively. Define

$$
\mu_n = \frac{p_1\tau(1-\tau)}{2h} E\left(\frac{1}{f_{y|u,x}(q_\tau(U_t, X^{(1)} - \Omega_{12}\Omega_{22}X^{(2)}))} \right)\nu_0,
$$

$$\sigma_n^2 = \frac{2(\tau(1-\tau))^2 p}{h} E\left(\frac{1}{f_{y|u,x}^2(q_\tau(U_t, X^{(1)} - \Omega_{12}\Omega_{22}X^{(2)}))}\right)$$

$$\int ((2K(t) - K * K(t))^2 dt,$$

$$T_2 = \sum_{t=1}^{n} \varepsilon_t(X_t^{(1)} - \Omega_{12}\Omega_{22}X_t^{(2)})\frac{h^2}{2}A^{(2)}(U_t)\mu_2 + o(nh^2),$$

$$T_4 = \frac{h^2\mu_2}{n}\sum_{t=1}^{n}\frac{f_{y|u,x}(q_\tau(U_t, X_t))A^{(2)}(U_t)}{f_u(U_t)\Omega_{11,2}}\sum_{j \neq t}\varepsilon_j(X_j^{(1)} - \Omega_{12}\Omega_{22}X_j^{(2)})K_h\left(\frac{U_j - U_t}{h}\right),$$

$$T_5 = \frac{1}{8}\sum_{t=1}^{n}f_{y|u,x}(q_\tau(U_t, X_t)(X_t^{(1)} - \Omega_{12}\Omega_{22}X_t^{(2)})(X_t^{(1)} - \Omega_{12}\Omega_{22}X_t^{(2)})^T h^4$$

$$(A^{(2)}(U_t)\mu_2)^2 + o(nh^4),$$

and

$$d_n^* = T_2 - T_4 - T_5.$$

Then, we have the asymptotic normality for the proposed test statistic stated in the following theorem.

Theorem 3.1: Suppose Assumptions A and B hold. Then, under H_0 in (3.3), as $nh^{3/2} \to \infty$ and $h \to 0$, we have

$$\sigma_n^{-1}(T_n - d_n^* - \mu_n) \to N(0, 1).$$

Proof: See Appendix.□

3.2.2   Test of Constancy of Partial Varying Coefficient

Now, we consider the hypothesis given in (3.2) with A_{10} being a vector of known constants. Following the similar approaches in the previous subsection, we can construct the quasi-likelihood function under null or alternative hypothesis by estimating the unknown functional coefficients using local linear technique with jackknife method. Then, by the same token, the generalized quasi-likelihood ratio (GQLR)

test statistic can be defined as

$$T_n = \ell(H_a) - \ell(H_0)$$
$$= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_1(U_t)X_t^{(1)} - \hat{A}_2(U_t)X_t^{(2)} \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t^{**} - \hat{A}_2^{**}(U_t)X_t^{(2)} \right),$$

where $\hat{A}_1(\cdot)$ and $\hat{A}_2(\cdot)$ are the same as in (3.4). $\hat{A}_2^{**}(\cdot)$ is is the nonparametric esti-
mator from the varying coefficient quantile model as $q_\tau(Y^{**}|U, X^{(2)}) = A_2^{**}(U)^T X^{(2)}$
by using the same estimation procedure as earlier, where $Y_t^{**} = Y_t - A_{10}X_t^{(1)}$; see
(2.6) for details. Define

$$T_2 = \sum_{t=1}^{n} \varepsilon_t(X_t^{(1)} - \Omega_{12}\Omega_{22}X_t^{(2)})\frac{h^2}{2}A_2^{(2)}(U_t)\mu_2 + o(nh^2),$$

$$T_4 = \frac{h^2\mu_2}{n}\sum_{t=1}^{n}\frac{f_{y|u,x}(q_\tau(U_t, X_t))A_2^{(2)}(U_t)}{f_u(U_t)\Omega_{11,2}}\sum_{j\neq t}\varepsilon_j(X_j^{(1)} - \Omega_{12}\Omega_{22}X_j^{(2)})K_h(\frac{U_j - U_t}{h}),$$

$$T_5 = \frac{1}{8}\sum_{t=1}^{n}f_{y|u,x}(q_\tau(U_t, X_t)(X_t^{(1)} - \Omega_{12}\Omega_{22}X_t^{(2)})(X_t^{(1)} - \Omega_{12}\Omega_{22}X_t^{(2)})^T$$
$$h^4(A_2^{(2)}(U_t)\mu_2)^2 + o(nh^4),$$

and

$$d_n = T_2 - T_4 - T_5.$$

Then, we have the following asymptotic result.

Theorem 3.2: Suppose Assumptions A and B hold and define μ_n and σ_n^2 as in
Theorem 3.1. Then, under H_0 in (3.2), as $nh^{3/2} \to \infty$ and $h \to 0$, we have

$$\sigma_n^{-1}(T_n - d_n - \mu_n) \to N(0, 1).$$

Proof: See Appendix.□

Remark 3.1: Here μ_n and σ_n are the same as μ_n and σ_n in Chapter 2 with p
replaced by p_1, X replaced by $X^{(1)} - \Omega_{12}\Omega_{22}X^{(2)}$, and and $\Omega^*(U_t)$ replaced by $\Omega_{11,2}$,

respectively.

3.2.3 Test of Constancy of Partial Varying Coefficient with Unknown Value

Finally, we consider the hypothesis given in (3.2) with A_{10} being a vector of unknown constants. To estimate the coefficient functions, under the alternative, we can continue to use the local linear technique with jackknife method. Under the null hypothesis, however, we need to use a new procedure to estimate constant and functional coefficients. Indeed, under the null hypothesis, the model becomes partially varying coefficient quantile regression model as follows

$$q_\tau(U_t, X_t) = \beta_\tau^T X_{t1} + \alpha_\tau(U_t)^T X_{t2}, \qquad (3.5)$$

which is similar to the model in Cai and Xiao (2012). To estimate the parameter β_τ and the function $\alpha_\tau(\cdot)$ in the above model, they proposed a consistent semi-parametric estimation procedure, described briefly as follows: First, β_τ is regarded as a function of U_t so that the model becomes a fully varying coefficient model and all coefficient functions can be estimated by a nonparametric fitting scheme; see (2.6) for details. Secondly, a root-n consistent estimator for β_τ is obtained by using the average method, denoted by $\hat{\beta}_\tau$. Finally, a nonparametric approach is applied to estimating $\alpha_\tau(\cdot)$, denoted by $\hat{\alpha}_\tau(\cdot)$, based on the partial quantile residual $Y_t^* = Y_t - \hat{\beta}_\tau^T X_{t1}$, where $\hat{\beta}_\tau$ is a root-n consistent estimator of β_τ. Under certain mild assumptions, Cai and Xiao (2012) derived the following asymptotic results

$$\sqrt{n}\left[\hat{\beta}_\tau - \beta_\tau\right] \to N(0, \Sigma_\beta)$$

for some Σ_β, and

$$\sqrt{nh_1}\left[\hat{\alpha}_\tau(u_0) - \alpha_\tau(u_0) - \frac{h_1^2 \mu_2 \alpha_\tau''(u_0)}{2} + o_p(h_1^2)\right] \to N(0, \Sigma_\alpha)$$

for some Σ_α. These results imply that the estimator for the parametric part is root-n consistent and the nonparametric estimate for nonparametric part has the regular nonparametric convergence rate.

Therefore, under the null hypothesis, the quasi-likelihood is given by

$$\ell(H_0) = \sum_{t=1}^{n} \rho_\tau \left(Y_t - \tilde{A}_1 X_t^{(1)} - \tilde{A}_2(U_t) X_t^{(2)} \right),$$

where $\tilde{A}_1 = \hat{\beta}_\tau$ and $\tilde{A}_2(\cdot) = \hat{\alpha}_\tau(\cdot)$ are the estimators discussed in (3.5). Then, the generalized quasi-likelihood ratio (GQLR) test is defined as follows:

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_1(U_t) X_t^{(1)} - \hat{A}_2(U_t) X_t^{(2)} \right) - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \tilde{A}_1 X_t^{(1)} - \tilde{A}_2(U_t) X_t^{(2)} \right)
\end{aligned}
$$

where $\hat{A}_1(\cdot)$ and $\hat{A}_2(\cdot)$ are the same as in (3.4). Thus, by using the same notation as Theorems 3.1 and 3.2, we have the following asymptotic result for the test statistic.

Theorem 3.3: Suppose Assumptions A and B hold. Then, under H_0, as $nh^{3/2} \to \infty$ and $h \to 0$, we have

$$\sigma_n^{-1}(T_n - d_n - \mu_n) \to N(0, 1).$$

Proof: See Appendix.□

3.3 The power of Test statistic

In this section, we consider the power of the quasi-likelihood ratio test under local alternative of the form

$$H_a : A_1(u) = A_{10} + \frac{1}{\sqrt{nh}} \Delta(u),$$

where $\Delta(u) = (\Delta_1(u), \Delta_2(u), ..., \Delta_p(u))^T$ is a vector of functions, satisfying $E(||\Delta(u)||_2^2) < \infty$ and A_{10} is a known constant vector. Define

$$d_{2n} = \frac{1}{2h} E(f_{y|u,x}(q_\tau|X_t, U_t)\Delta(U_t)^T X_t^{(1)}(X_t^{(1)})^T \Delta(U_t))$$

and $\sigma_n^* = \sqrt{\sigma_n^2 + (\tau(1-\tau))E\Delta^T(U)X^{(1)}(X^{(1)})^T\Delta(U)}$. Then, we have the asymptotical distribution of test statistic T_n.

Theorem 3.4: Assume the same conditions as in Theorems 3.1and 3.2 hold. Then, under H_a, We have

$$\sigma_n^{*-1}(T_n - \mu_n - d_n + d_{2n}) \to N(0,1).$$

Proof: See Appendix.□

3.4   Simulation Studies

Example 3.1: In this simulated example, I consider the following data generating process:

$$Y_t = a_1(U_t)X_{1t} + a_2(U_t)X_{2t} + e_t, \quad 1 \le t \le n, \tag{3.6}$$

where $a_1(u) = 2$, $a_2(u) = \cos(\sqrt{2}\pi u)$, U_t is generated from uniform $(0, 3)$ independently, $e_t \sim N(0, 0.3)$, $X_{1t} \sim N(0.5, 0.4)$ and $X_{2t} \sim N(0.75, 0.4)$. Then, the corresponding quantile regression is

$$q_\tau(U_t, X_t) = a_{0,\tau}(U_t) + a_{1,\tau}(U_t)X_{t1} + a_{2,\tau}(U_t)X_{t2},$$

where $a_{0,\tau}(U_t) = \sqrt{0.3}\Phi^{-1}(\tau)$, $a_{1,\tau}(u) = a_1(u)$, $a_{2,\tau}(u) = a_2(u)$, and $\Phi^{-1}(\tau)$ is the τ-th quantile of $N(0,1)$. The purpose of this simulated example is to exam the finite sample performance of the proposed test for testing if $a_1(u)$ is equal to a known constant.

I choose the sample sizes as $n = 250$, 500 and 800 and repeat the simulation $m = 1000$ times. I apply the jackknife method and local linear fitting technique to estimate the coefficient and select the bandwidth as discussed in Section 3. I report the different testing nominal sizes at 1%, 5% and 10% for different quantiles $\tau = 0.2$, 0.4, 0.6 and 0.8. The simulated test sizes for different settings are listed in Table 3.1. One can observe from Table 3.1 that the empirical test sizes are very close to the true nominal sizes for all settings. This means that our proposed test can give the correct test size.

Table 3.1: : Finite sample rejection rates for example 3.1

nominal size	sample size	$\tau = 0.2$	$\tau = 0.4$	$\tau = 0.6$	$\tau = 0.8$
	$n = 250$	0.116	0.109	0.108	0.103
10%	$n = 500$	0.104	0.109	0.105	0.098
	$n = 800$	0.112	0.115	0.105	0.099
	$n = 250$	0.068	0.057	0.053	0.050
5%	$n = 500$	0.050	0.055	0.055	0.048
	$n = 800$	0.058	0.061	0.049	0.053
	$n = 250$	0.020	0.013	0.019	0.013
1%	$n = 500$	0.009	0.017	0.017	0.014
	$n = 800$	0.013	0.014	0.014	0.013

To demonstrate the power of the proposed test, the power function is evaluated under a sequence of the alternative models indexed by λ as

$$H_a : a_{1,\tau}(u) = 2 + \lambda \frac{1}{\sqrt{nh}} \Delta(u), \quad 0 \leq \lambda \leq 1, \tag{3.7}$$

where $\Delta(u) = 2u^4 e^{-u/10}$ and $\lambda = 0.05\,i$ for $1 \leq i \leq 20$. Given the significance level 5%, the power function $p(\lambda)$ is estimated based on the relative frequency of T_n over the critical value among 1000 simulations. We plot the power curves in Figure 3.1 for all settings, from which, we can find that the proposed test statistic is powerful. Example 3.2: In this simulated example, all settings are the same as those in Example 3.1 but the true functional form for $a_1(u)$ is considered as unknown when

constructing the quasi-likelihood under the null hypothesis. The simulation results for test sizes are reported in Table 3.2, from which one can observe that the performance of the proposed test is reasonably well. Similarly, we compute the power

Table 3.2: : Finite sample rejection rates for example 3.2

nominal size	sample size	$\tau = 0.2$	$\tau = 0.4$	$\tau = 0.6$	$\tau = 0.8$
	$n = 250$	0.105	0.092	0.108	0.099
10%	$n = 500$	0.098	0.103	0.115	0.094
	$n = 800$	0.099	0.100	0.106	0.089
	$n = 250$	0.049	0.045	0.059	0.055
5%	$n = 500$	0.046	0.056	0.057	0.049
	$n = 800$	0.049	0.049	0.054	0.052
	$n = 250$	0.017	0.009	0.014	0.010
1%	$n = 500$	0.010	0.010	0.015	0.010
	$n = 800$	0.010	0.016	0.016	0.013

curves which are displayed in Figure 3.2, from which we can find that the proposed test statistic is powerful.

Example 3.3: In this simulated example, I consider the simplest model

$$Y_t = a_1(U_t)X_{1t} + a_2(U_t)X_{2t} + e_t, \quad 1 \leq t \leq n, \tag{3.8}$$

where $a_1(u) = \sin(\sqrt{2}\pi u)$, $a_2(u) = \cos(\sqrt{2}\pi u)$, U_t is generated from uniform $(0,3)$ independently, $e_t \sim N(0, 0.3)$, $X_{1t} \sim N(0.5, 0.4)$ and $X_{2t} \sim N(0.75, 0.4)$. Then, the corresponding quantile regression is

$$q_\tau(U_t, X_t) = a_{0,\tau}(U_t) + a_{1,\tau}(U_t)X_{1t} + a_{2,\tau}(U_t)X_{2t},$$

where $a_{0,\tau}(u) = \sqrt{0.3}\Phi^{-1}(\tau)$, $a_{1,\tau}(u) = a_1(u)$, $a_{2,\tau}(u) = a_2(u)$, and $\Phi^{-1}(\tau)$ is the τ-th quantile of the $N(0,1)$. Other settings are the same as those in Example 3.1. The simulated sizes are presented in Table 3.3. The power function is evaluated

Table 3.3: : Finite sample rejection rates for example 3.3

nominal size	sample size	$\tau = 0.2$	$\tau = 0.4$	$\tau = 0.6$	$\tau = 0.8$
	$n = 250$	0.123	0.105	0.098	0.124
10%	$n = 500$	0.122	0.096	0.105	0.110
	$n = 800$	0.109	0.101	0.114	0.106
	$n = 250$	0.062	0.072	0.052	0.058
5%	$n = 500$	0.058	0.050	0.053	0.064
	$n = 800$	0.055	0.054	0.058	0.048
	$n = 250$	0.020	0.032	0.018	0.013
1%	$n = 500$	0.018	0.017	0.018	0.022
	$n = 800$	0.017	0.013	0.024	0.013

under a sequence of the alternative models indexed by $0 \leq \lambda \leq 1$,

$$H_a : a_{1,\tau}(u) = \sin(\sqrt{2}\pi\, u) + \frac{\lambda}{\sqrt{nh}}\Delta(u),$$

where $\Delta(u) = u^4 e^{-u/10}$ and $\lambda = 0.05\, i$ for $1 \leq i \leq 20$. Given the significance level 5%, we compute the power curves as functions of λ and the simulated power curves are given in Figure 3.3. The conclusions similar to Examples 3.1 and 3.2 can be made.

3.5 A Real Example

In this section, I continue to consider the application of the proposed methodologies to study the Boston house price data. According to the discussion in Section 2.6, the coefficients for X_3 and X_4 may be constant if they are included in the model. Therefore, a semiparametric model is appropriate if the model includes these two variables. Indeed, Xu (2005) considered the following model

$$q_\tau(U_t, X_t) = a_{0,\tau}(U_t) + a_{1,\tau}(U_t)X_{t1} + a_{2,\tau}(U_t)X_{t2}^* + a_{4,\tau}X_{t4}. \tag{3.9}$$

To see if the above model in (3.9) is correctly specified, we consider the following varying coefficient quantile model

$$q_\tau(U_t, X_t) = a_{0,\tau}(U_t) + a_{1,\tau}(U_t)X_{t1} + a_{2,\tau}(U_t)X_{t2}^* + a_{4,\tau}(U_t)X_{t4} \qquad (3.10)$$

and then test the null hypothesis $H_0 : a_{4,\tau}(u) = a_{4,\tau,0}$, where $a_{4,\tau,0}$ is an unknown parameter. That is to test model (3.10) against model (3.9). Thus, we can use our test procedure proposed in Section 3.2.3 for this testing purpose. I calculate the quasi-likelihood using semiparametric quantile regression method proposed in Cai and Xiao (2012) under the null hypothesis and calculate the quasi-likelihood using local linear fitting method with jackknife technique. The p-values for several quantiles are reported in Table 3.4. Therefore, one can see from Table 3.4 that the coefficient for X_{t4} is a constant for all quantiles. This implies that model (3.9) is appropriate to analyze the Boston house pricing data.

Table 3.4: : The p-values for testing model (3.10) versus model (3.9)

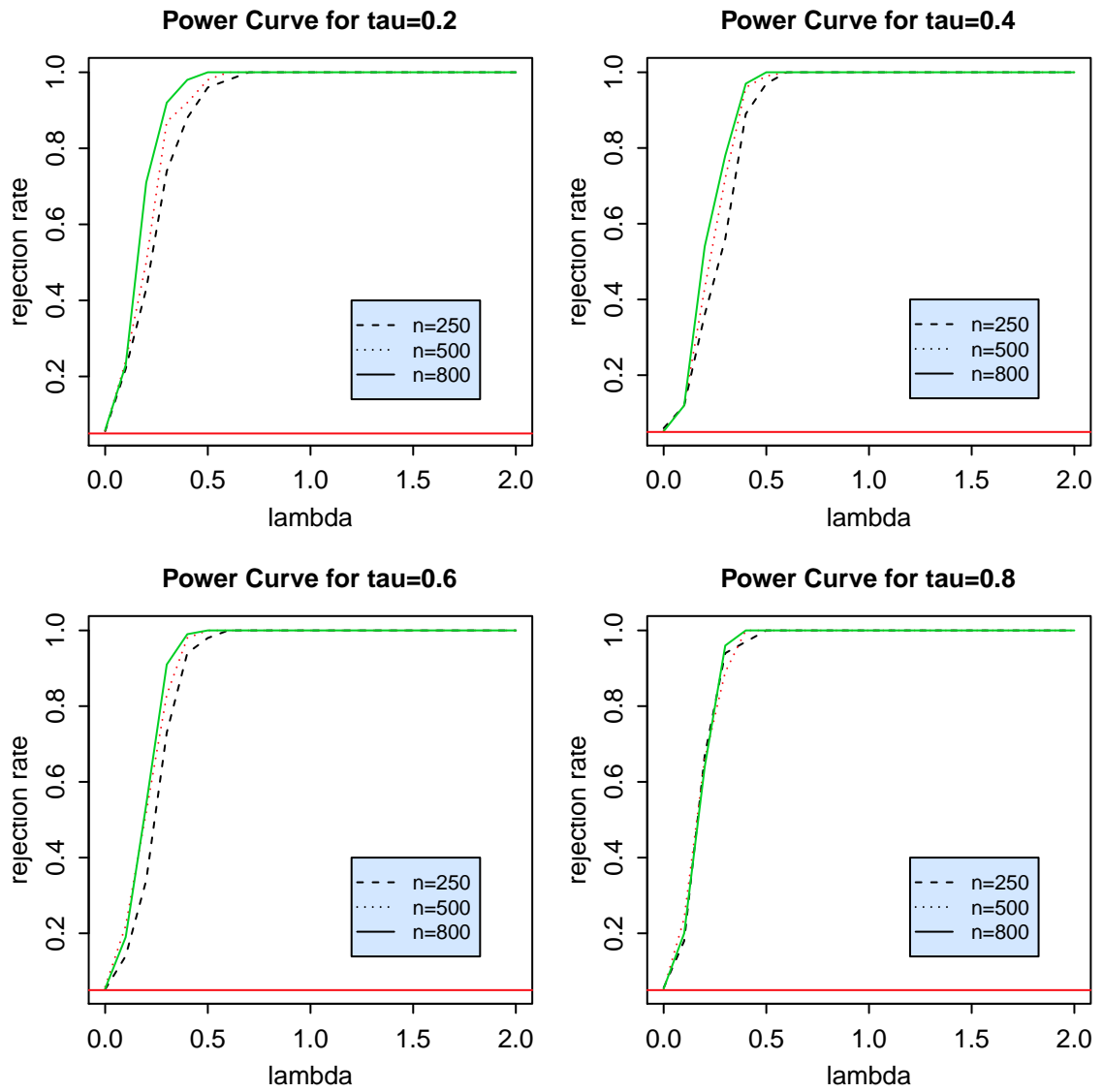| τ	0.2	0.4	0.6	0.8
p-value	0.880	0.313	0.610	0.532

Figure 3.1: The plot of power curves against λ for the testing hypothesis in Example 3.1 with nominal test size 5%. The dashed line is for $n = 400$, the solid line is for $n = 600$ and the dashed-dotted line is for $n = 800$.

Figure 3.2: The plot of power curves against λ for the testing hypothesis in Example 3.2 with nominal test size 5%. The dashed line is for $n = 400$, the solid line is for $n = 600$ and the dashed-dotted line is for $n = 800$.
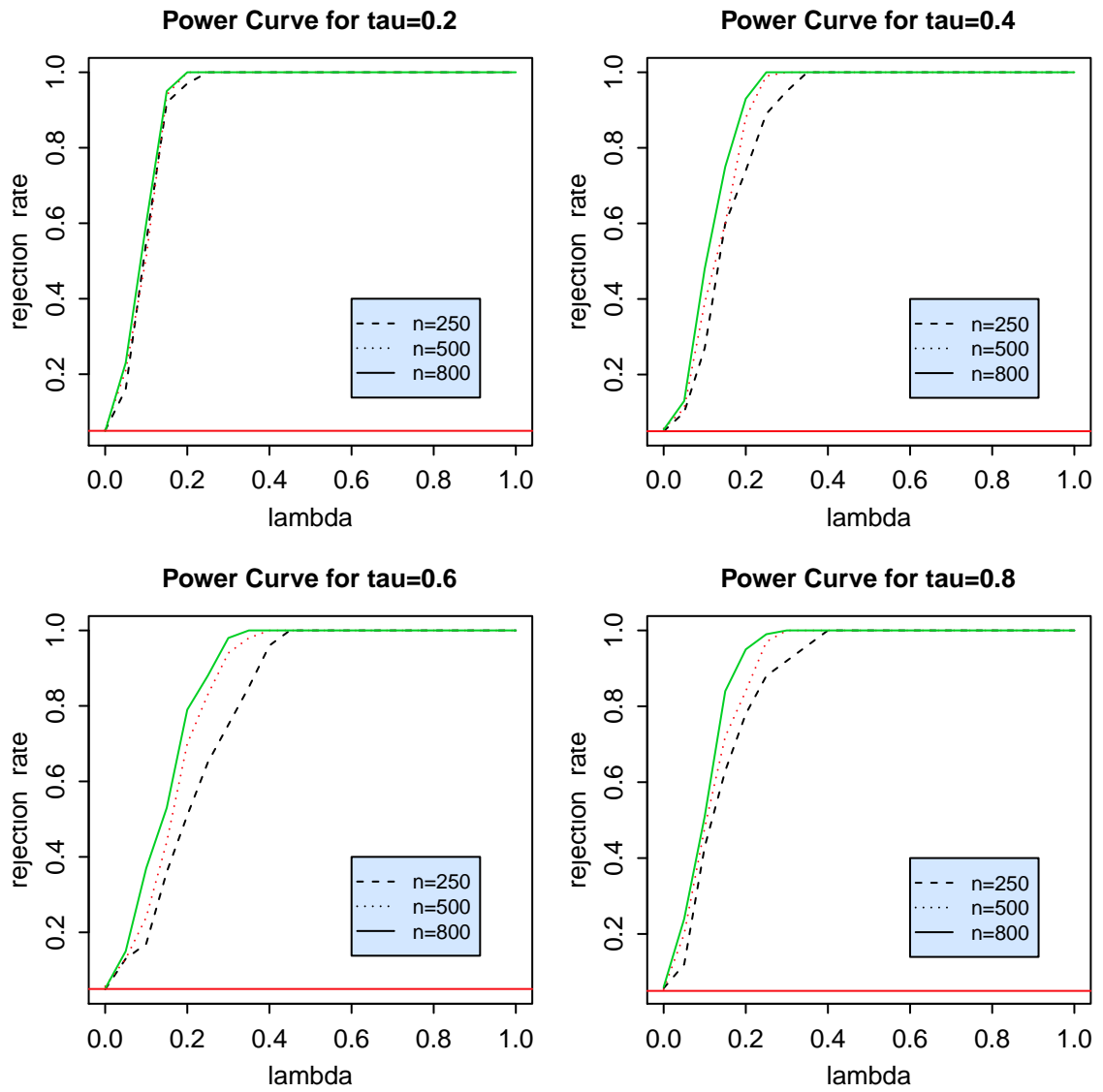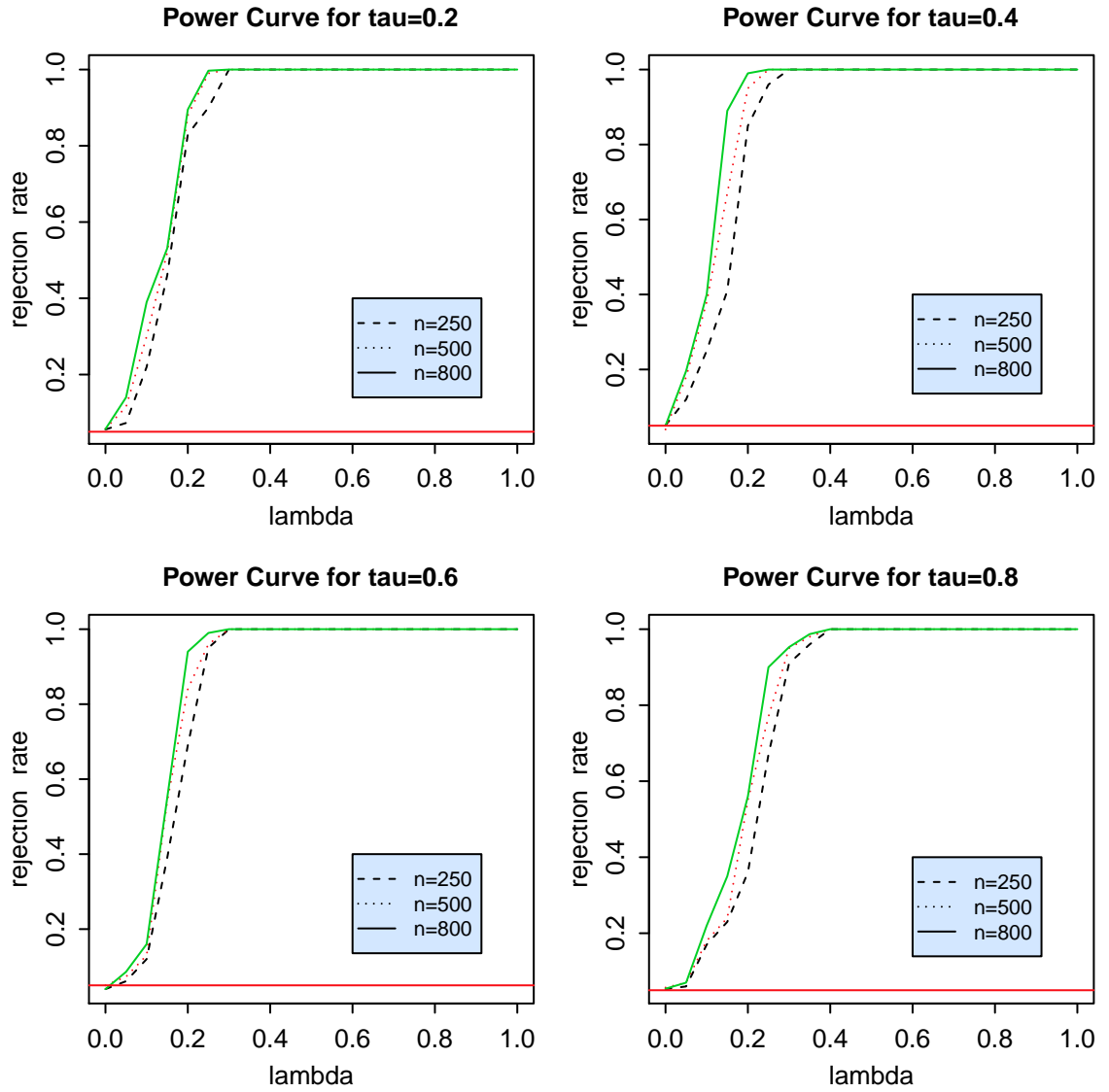
Figure 3.3: The plot of power curves against λ for the testing hypothesis in Example 3.3 with nominal test size 5%. The dashed line is for $n = 400$, the solid line is for $n = 600$ and the dashed-dotted line is for $n = 800$.

3.6   Complements

In this section, I will give the derivations of the main results presented in pre-vious sections of this chapter. Before moving to the detailed proofs, I will state the following useful lemma first.

Lemma 3.1:   Analogously to the arguments for $\hat{A}$ we have the following result.

$$
\begin{aligned}
\hat{A}_{-2t}(U_t) - A_2(U_t) &= \frac{(\Omega_{22}(U_t))^{-1}}{(\sqrt{nh})^2 f_u(U_t)} \sum_{i \neq t}^{n} \varphi_\tau(Y_i^*) X_i^{(2)} K(\frac{U_i - U_t}{h}) + o_p(\frac{1}{\sqrt{nh}}) \\
&= \frac{(\Omega_{22}(U_t))^{-1}}{(\sqrt{nh})^2 f_u(U_t)} \sum_{i \neq t}^{n} \varepsilon_i X_i^{(2)} K(\frac{U_i - U_t}{h}) \\
&\quad \frac{(\Omega_{22}(U_t))^{-1}}{(\sqrt{nh})^2 f_u(U_t)} \sum_{i \neq t}^{n} (\varphi_\tau(Y_i^*) - \varepsilon_i) X_i^{(2)} K(\frac{U_i - U_t}{h}) + o_p(\frac{1}{\sqrt{nh}}) \\
&\equiv R_1^* + R_2^*,
\end{aligned}
$$

where $Y_i^* = Y_i - A_{10} X_i^{(1)} - A_2 X_i^{(2)} - A_2' X_i^{(2)} (U_k - U_i)$.

Proof of Lemma 3.1

From the derivation of Theorem 2.1, we have

$$
\left(\hat{A}_{-t}(u) - A(u) \right)
$$

$$
\begin{aligned}
&= \frac{1}{(\sqrt{nh})^2 f_u(U_t)} \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} \sum_{i \neq t}^{n} \varphi_\tau(Y_i^*) \begin{pmatrix} X_i^{(1)} \\ X_i^{(2)} \end{pmatrix} \\
&\quad K(\frac{U_i - U_t}{h}) + o_p(\frac{1}{\sqrt{nh}}) \\
&= \sum_{i \neq t}^{n} \frac{1}{(\sqrt{nh})^2 f_u(U_t)} \varphi_\tau(Y_i^*) \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} \begin{pmatrix} X_i^{(1)} \\ X_i^{(2)} \end{pmatrix} K(\frac{U_i - U_t}{h}).
\end{aligned}
$$

We use $Y_i^* = Y_i - A_{10}X_i^{(1)}$ as new response variable, by following the similar procedure to the derivation of Theorem 2.1, to obtain

$$
\begin{aligned}
\hat{A}_{-2t}(u) - A_2(u) &= \sum_{i \neq t}^{n} \Omega_{22}^{-1} X_i^{(2)} \times \frac{1}{(\sqrt{nh})^2 f_u(U_t)} \varphi_\tau(Y_i^*) K(\frac{U_i - U_t}{h}) \\
&= R_1^* + R_2^*.
\end{aligned}
$$

That finishes the proof for lemma 3.1.

Lemma 3.2: (Theorem 1 in Cai and Xiao (2012)) Under some assumptions,

$$
\sqrt{n}[\hat{\beta}_\tau - \beta_\tau - B_\beta] \to N(0, \Sigma_\beta),
$$

where the asymptotic bias term is $B_\beta = h^2 \mu_2 (B_1^* - B_2^*/2)$, and the asymptotic variance

$$
\begin{aligned}
\Sigma_\beta &= \tau(1 - \tau) E[e_1^T (\Omega^*(U_1))^{-1} \Omega(U_1)(\Omega^*(U_1))^{-1} e_1] \\
&\quad + 2 \sum_{s=1}^{\infty} cov((e_1^T \Omega^*(U_1))^{-1}) X_1 \eta_1, (e_1^T \Omega^*(U_{s+1}))^{-1}) X_{s+1} \eta_{s+1}).
\end{aligned}
$$

Lemma 3.3: (Theorem 2 in Cai and Xiao (2012)) Under some assumptions, the local linear estimator of $\alpha(u_0)$ has the following asymptotic distribution:

$$
\sqrt{nh_1}[\hat{\alpha}_\tau(u_0) - \alpha_\tau(u_0) - \frac{h_1^2 \mu_2 \alpha''(u_0)}{2} + o_p(h_1^2)] \to N(0, \Sigma_\alpha),
$$

where $\Sigma_\alpha = \tau(1 - \tau)v_0 \Sigma_\alpha(u_0)/f_u(u_0)$.

Proof of Theorem 3.1: Note that

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau\left(Y_t - \hat{A}_1(u)X_k^{(1)} - \hat{A}_2(u)X_k^{(2)}\right) - \sum_{t=1}^{n} \rho_\tau\left(Y_t - A_{10}X_k^{(1)} - \hat{A}_2(u)X_k^{(2)}\right) \\
&= \sum_{t=1}^{n} \rho_\tau\left(Y_t - \hat{A}_1(u)X_k^{(1)} - \hat{A}_2(u)X_k^{(2)}\right) - \sum_{t=1}^{n} \rho_\tau\left(Y_t - A_{10}X_k^{(1)} - A_{20}(u)X_k^{(2)}\right) \\
&\quad + \sum_{t=1}^{n} \rho_\tau\left(Y_t - A_{10}X_k^{(1)} - A_{20}(u)X_k^{(2)}\right) - \sum_{t=1}^{n} \rho_\tau\left(Y_t - A_{10}X_k^{(1)} - \hat{A}_2(u)X_k^{(2)}\right).
\end{aligned}
$$

Follow the similar derivation to Theorem 2.2, one can get

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= -\sum_{t=1}^{n}(\hat{A}_{-t}(U_t) - A(U_t))^T X_t \varepsilon_t \\
&\quad + \frac{1}{2}\sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t))(\hat{A}_{-t}(U_t) - A(U_t))^T X_t X_t^T(\hat{A}_{-t}(U_t) - A(U_t)) \\
&\quad + \sum_{t=1}^{n}(\hat{A}_{-2t}(U_t) - A_2(U_t))^T X_t^{(2)}\varepsilon_t - \frac{1}{2}\sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t)) \\
&\quad (\hat{A}_{-2t}(U_t) - A_2(U_t))^T X_t^{(2)}(X_t^{(2)})^T(\hat{A}_{-2t}(U_t) - A_2(U_t)) \\
&= \sum_{t=1}^{n}[(\hat{A}_{-2t}(U_t) - A_2(U_t))^T X_t^{(2)} - (\hat{A}_{-t}(U_t) - A(U_t))^T X_t]\varepsilon_t \\
&\quad + \frac{1}{2}\sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t))[(\hat{A}_{-t}(U_t) - A(U_t))^T X_t X_t^T(\hat{A}_{-t}(U_t) - A(U_t)) \\
&\quad - (\hat{A}_{-2t}(U_t) - A_2(U_t))^T X_t^{(2)}(X_t^{(2)})^T \\
&\quad (\hat{A}_{-2t}(U_t) - A_2(U_t))] \\
&= \sum_{t=1}^{n} D_1\varepsilon_t + \frac{1}{2}\sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t))D_2,
\end{aligned}
$$

where $D_1 = (\hat{A}_{-2t}(U_t) - A_2(U_t))^T X_t^{(2)} - (\hat{A}_{-t}(U_t) - A(U_t))^T X_t$ and

$$
\begin{aligned}
D_2 &= (\hat{A}_{-t}(U_t) - A(U_t))^T X_t X_t^T (\hat{A}_{-t}(U_t) - A(U_t)) \\
&\quad - (\hat{A}_{-2t}(U_t) - A_2(U_t))^T X_t^{(2)} (X_t^{(2)})^T (\hat{A}_{-2t}(U_t) - A_2(U_t)),
\end{aligned}
$$

For D_1, apply Theorem 2.1 and Lemma 3.1 to get

$$
\begin{aligned}
D_1 &= (\hat{A}_{-2t}(U_t) - A_2(U_t))^T X_t^{(2)} - (\hat{A}_{-t}(U_t) - A(U_t))^T X_t \\
&= \frac{(\Omega_{22}(U_t))^{-1}}{(\sqrt{nh})^2 f_u(U_t)} \sum_{i \neq t}^n \varphi_\tau(Y_i^*) X_i^{(2)} K\left(\frac{U_i - U_t}{h}\right) X_t^{(2)} \\
&\quad - \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} \frac{1}{nh f_u(u_0)} \sum_{t=1}^n \varphi_\tau(Y_i^*) X_i K\left(\frac{U_t - u_0}{h}\right) X_t \\
&= \frac{1}{(\sqrt{nh})^2 f_u(U_t)} \sum_{i \neq t}^n \varphi_\tau(Y_i^*) [X_i^{(2)} (\Omega_{22}(U_t))^{-1} X_t^{(2)} - X_i \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} \\
&\quad X_t] K\left(\frac{U_i - U_t}{h}\right) \\
&= \frac{1}{(\sqrt{nh})^2 f_u(U_t)} \sum_{i \neq t}^n \varphi_\tau(Y_i^*) D_{11} K\left(\frac{U_i - U_t}{h}\right).
\end{aligned}
$$

Then,

$$
\begin{aligned}
D_{11} &= X_i^{(2)} (\Omega_{22}(U_t))^{-1} X_t^{(2)} - X_i \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} X_t \\
&= X_i^{(2)} (\Omega_{22}(U_t))^{-1} X_t^{(2)} - \begin{pmatrix} X_i^{(1)} & X_i^{(2)} \end{pmatrix} \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} \begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \end{pmatrix} \\
&= X_i^{(2)} (\Omega_{22}(U_t))^{-1} X_t^{(2)} - X_i^{(1)} (\Omega_{11} - \Omega_{12} \Omega_{22}^{-1} \Omega_{21})^{-1} X_t^{(1)} \\
&\quad + X_i^{(1)} (\Omega_{11} - \Omega_{12} \Omega_{22}^{-1} \Omega_{21})^{-1} \Omega_{12} \Omega_{22}^{-1} X_t^{(2)} - X_i^{(2)} \Omega_{22}^{-1} \Omega_{21} (\Omega_{11} - \Omega_{12} \Omega_{22}^{-1} \\
&\quad \Omega_{21})^{-1} X_t^{(1)} - X_i^{(2)} (\Omega_{22}(U_t))^{-1} X_t^{(2)} - X_i^{(2)} \Omega_{22}^{-1} \Omega_{21} (\Omega_{11} - \Omega_{12} \Omega_{22}^{-1}
\end{aligned}
$$

$$\Omega_{21})^{-1}\Omega_{12}\Omega_{22}^{-1}X_t^{(2)}$$
$$= -(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})$$
$$= -(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}).$$

Therefore,

$$
\begin{aligned}
D_1 &= -\frac{1}{(\sqrt{nh})^2 f_u(U_t)}\sum_{i\neq t}^n \varphi_\tau(Y_i^*)(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})K(\frac{U_i - U_t}{h}) \\
&= -\frac{1}{(\sqrt{nh})^2 f_u(U_t)}\sum_{i\neq t}^n \varepsilon_i(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})K(\frac{U_i - U_t}{h}) \\
&\quad -\frac{1}{(\sqrt{nh})^2 f_u(U_t)}\sum_{i\neq t}^n (\varphi_\tau(Y_i^*) - \varepsilon_i)(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})K(\frac{U_i - U_t}{h}).
\end{aligned}
$$

Next, for D_2

$$
\begin{aligned}
D_2 &= (\hat{A}_{-t}(U_t) - A(U_t))^T X_t X_t^T (\hat{A}_{-t}(U_t) - A(U_t)) \\
&\quad - (\hat{A}_{-2t}(U_t) - A_2(U_t))^T X_t^{(2)}(X_t^{(2)})^T (\hat{A}_{-2t}(U_t) - A_2(U_t)) \\
&= \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} \frac{1}{nh f_u(u_0)}\sum_{t=1}^n \varphi_\tau(Y_i^*)X_i K(\frac{U_t - u_0}{h})X_t X_t^T \\
&\quad \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} \frac{1}{nh f_u(u_0)}\sum_{t=1}^n \varphi_\tau(Y_i^*)X_i K(\frac{U_t - u_0}{h}) \\
&\quad - \frac{(\Omega_{22}(U_t))^{-1}}{(\sqrt{nh})^2 f_u(U_t)}\sum_{i\neq t}^n \varphi_\tau(Y_i^*)X_i^{(2)} K(\frac{U_i - U_t}{h})X_t^{(2)}(X_t^{(2)})^T \\
&\quad \frac{(\Omega_{22}(U_t))^{-1}}{(\sqrt{nh})^2 f_u(U_t)}\sum_{i\neq t}^n \varphi_\tau(Y_i^*)X_i^{(2)} K(\frac{U_i - U_t}{h})
\end{aligned}
$$

$$= \quad (\frac{1}{nhf_u(u_0)})^2 \sum_{t=1}^{n} \sum_{t=1}^{n} \varphi_\tau(Y_i^*)\varphi_\tau(Y_i^*)X_i \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} X_t X_t^T$$

$$\begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} X_i K(\frac{U_t - u_0}{h})K(\frac{U_t - u_0}{h})$$

$$- (\frac{1}{nhf_u(u_0)})^2 \sum_{i \neq t}^{n} \sum_{i \neq t}^{n} \varphi_\tau(Y_i^*)\varphi_\tau(Y_i^*)X_i^{(2)}(\Omega_{22}(U_t))^{-1}X_t^{(2)}(X_t^{(2)})^T(\Omega_{22}(U_t))^{-1}$$

$$X_i^{(2)} K(\frac{U_i - U_t}{h})K(\frac{U_i - U_t}{h})$$

$$= \quad (\frac{1}{nhf_u(u_0)})^2 \sum_{t=1}^{n} \sum_{t=1}^{n} \varphi_\tau(Y_i^*)\varphi_\tau(Y_i^*)D_{21}K(\frac{U_t - u_0}{h})K(\frac{U_t - u_0}{h})$$

$$- (\frac{1}{nhf_u(u_0)})^2 \sum_{i \neq t}^{n} \sum_{i \neq t}^{n} \varphi_\tau(Y_i^*)\varphi_\tau(Y_i^*)D_{22}K(\frac{U_i - U_t}{h})K(\frac{U_i - U_t}{h}),$$

where

$$D_{21} \quad = \quad X_i \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} X_t X_t^T \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} X_i,$$

$$D_{22} \quad = \quad X_i^{(2)}(\Omega_{22}(U_t))^{-1}X_t^{(2)}(X_t^{(2)})^T(\Omega_{22}(U_t))^{-1}X_i^{(2)}.$$

For D_{21},

$$\begin{aligned}
D_{21} \quad = \quad & [X_i^{(1)}(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})^{-1}X_t^{(1)} - X_i^{(1)}(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})^{-1}\Omega_{12}\Omega_{22}^{-1}X_t^{(2)} \\
& + X_i^{(2)}\Omega_{22}^{-1}\Omega_{21}(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})^{-1}X_t^{(1)} + X_i^{(2)}(\Omega_{22}(U_t))^{-1}X_t^{(2)} \\
& + X_i^{(2)}\Omega_{22}^{-1}\Omega_{21}(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})^{-1}\Omega_{12}\Omega_{22}^{-1}X_t^{(2)}] \\
& \times [X_t^{(1)}(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})^{-1}X_j^{(1)} - X_t^{(1)}(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})^{-1}\Omega_{12}\Omega_{22}^{-1}X_j^{(2)} \\
& + X_t^{(2)}\Omega_{22}^{-1}\Omega_{21}(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})^{-1}X_j^{(1)} + X_t^{(2)}(\Omega_{22}(U_t))^{-1}X_j^{(2)}
\end{aligned}$$

$$+ X_t^{(2)}\Omega_{22}^{-1}\Omega_{21}(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})^{-1}\Omega_{12}\Omega_{22}^{-1}X_j^{(2)}]$$

$$= (X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})$$

$$\times (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1})$$

$$+ (X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})$$

$$X_t^{(2)}(\Omega_{22}(U_t))^{-1}X_j^{(2)} + X_i^{(2)}(\Omega_{22}(U_t))^{-1}X_t^{(2)}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}$$

$$(X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1}) + X_i^{(2)}(\Omega_{22}(U_t))^{-1}X_t^{(2)}(X_t^{(2)})^T(\Omega_{22}(U_t))^{-1}X_i^{(2)}.$$

Thus, we get

$$
\begin{aligned}
D_2 &= (\frac{1}{nhf_u(u_0)})^2\sum_{t=1}^n\sum_{t=1}^n\varphi_\tau(Y_i^*)\varphi_\tau(Y_i^*)[(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})\times(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_j^{(1)} - X_j^{(2)}\Omega_{12} \\
&\quad \Omega_{22}^{-1}) + (X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \\
&\quad X_t^{(2)}(\Omega_{22}(U_t))^{-1}X_j^{(2)} + X_i^{(2)}(\Omega_{22}(U_t))^{-1}X_t^{(2)}(X_t^{(1)} - X_t^{(2)} \\
&\quad \Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1})]K(\frac{U_t - u_0}{h})K(\frac{U_t - u_0}{h}) \\
&= (\frac{1}{nhf_u(u_0)})^2\sum_{t=1}^n\sum_{t=1}^n(\varphi_\tau(Y_i^*) - \varepsilon_i + \varepsilon_i)(\varphi_\tau(Y_i^*) - \varepsilon_j + \varepsilon_j)[(X_i^{(1)} - \\
&\quad X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)} \\
&\quad \Omega_{12}\Omega_{22}^{-1})\times(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1}) \\
&\quad + (X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \\
&\quad X_t^{(2)}(\Omega_{22}(U_t))^{-1}X_j^{(2)} \\
&\quad + X_i^{(2)}(\Omega_{22}(U_t))^{-1}X_t^{(2)}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1})] \\
&\quad K(\frac{U_t - u_0}{h})K(\frac{U_t - u_0}{h}) \\
&= (\frac{1}{nhf_u(u_0)})^2\sum_{t=1}^n\sum_{t=1}^n\varepsilon_i\varepsilon_j[(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})\times(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1}) \\
&\quad K(\frac{U_t - u_0}{h})K(\frac{U_t - u_0}{h}) + D_{23} + D_{24},
\end{aligned}
$$

where

$$
\begin{aligned}
D_{23} &= 2(\frac{2}{nhf_u(u_0)})^2 \sum_{t=1}^{n}\sum_{t=1}^{n} \varepsilon_i(\varphi_\tau(Y_i^*) - \varepsilon_j)[(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1}) \\
&\quad (\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \times (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1}) + (X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \times X_t^{(2)}(\Omega_{22}(U_t))^{-1}X_j^{(2)},
\end{aligned}
$$

and

$$
\begin{aligned}
D_{24} &= (\frac{1}{nhf_u(u_0)})^2 \sum_{t=1}^{n}\sum_{t=1}^{n} (\varphi_\tau(Y_i^*) - \varepsilon_i)(\varphi_\tau(Y_i^*) - \varepsilon_j)[(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1}) \\
&\quad (\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \times (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1}) + (X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)} \\
&\quad \Omega_{12}\Omega_{22}^{-1}) \times X_t^{(2)}(\Omega_{22}(U_t))^{-1}X_j^{(2)}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= -\sum_{t=1}^{n} \frac{1}{(\sqrt{nh})^2 f_u(U_t)} \sum_{i\neq t}^{n} \varepsilon_i(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})K(\frac{U_i - U_t}{h})\varepsilon_t + \frac{1}{2}\sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t)) \\
&\quad (\frac{1}{nhf_u(u_0)})^2 \sum_{t=1}^{n}\sum_{t=1}^{n} \varepsilon_i\varepsilon_j[(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1})]K(\frac{U_t - u_0}{h})K(\frac{U_t - u_0}{h}) - T_2^* + T_4^* + T_5^* + T_6^* + T_7^* \\
&\quad + T_8^* + T_9^*,
\end{aligned}
$$

where

$$
\begin{aligned}
T_2^* &= \sum_{t=1}^{n} \frac{1}{(\sqrt{nh})^2 f_u(U_t)} \sum_{i \neq t}^{n} (\varphi_\tau(Y_i^*) - \varepsilon_j)(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})K(\frac{U_i - U_t}{h})\varepsilon_t, \\
T_4^* &= \sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t))(\frac{2}{nhf_u(u_0)})^2 \sum_{t=1}^{n}\sum_{t=1}^{n} \varepsilon_i(\varphi_\tau(Y_i^*) - \varepsilon_j) \\
&\quad (X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1}), \\
T_5^* &= \frac{1}{2}\sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t))(\frac{1}{nhf_u(u_0)})^2 \sum_{t=1}^{n}\sum_{t=1}^{n} (\varphi_\tau(Y_i^*) - \varepsilon_i)(\varphi_\tau(Y_i^*) - \varepsilon_j) \\
&\quad [(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \times (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \\
&\quad (\Omega_{11,2})^{-1}(X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1}), \\
T_6^* &= (\frac{1}{nhf_u(u_0)})^2 \sum_{t=1}^{n}\sum_{t=1}^{n} \varepsilon_i\varepsilon_j(X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \times X_t^{(2)}(\Omega_{22}(U_t))^{-1}X_j^{(2)}K(\frac{U_t - u_0}{h})K(\frac{U_t - u_0}{h}), \\
T_7^* &= (\frac{1}{nhf_u(u_0)})^2 \sum_{t=1}^{n}\sum_{t=1}^{n} \varepsilon_i\varepsilon_j(X_i^{(2)}(\Omega_{22}(U_t))^{-1}X_t^{(2)}(X_t^{(1)} - X_t^{(2)}\Omega_{12} \\
&\quad \Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1})K(\frac{U_t - u_0}{h})K(\frac{U_t - u_0}{h}), \\
T_8^* &= \sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t))(\frac{2}{nhf_u(u_0)})^2 \sum_{t=1}^{n}\sum_{t=1}^{n} \varepsilon_i(\varphi_\tau(Y_i^*) - \varepsilon_j)(X_i^{(1)} - \\
&\quad X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1}) \times X_t^{(2)}(\Omega_{22}(U_t))^{-1}X_j^{(2)},
\end{aligned}
$$

and

$$
\begin{aligned}
T_9^* &= \frac{1}{2}\sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t))(\frac{1}{nhf_u(u_0)})^2 \sum_{t=1}^{n}\sum_{t=1}^{n} (\varphi_\tau(Y_i^*) - \varepsilon_i)(\varphi_\tau(Y_i^*) - \varepsilon_j) \\
&\quad (X_i^{(1)} - X_i^{(2)}\Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} - X_t^{(2)}\Omega_{12}\Omega_{22}^{-1})X_t^{(2)}(\Omega_{22}(U_t))^{-1}X_j^{(2)}.
\end{aligned}
$$

One can show easily that as $nh^{3/2} \to \infty$,

$$
E(T_6^*)^2 = O(\frac{1}{n^2h^4}) = o(\frac{1}{h}) \quad \text{and} \quad E(T_7^*)^2 = O(\frac{1}{n^2h^4}) = o(\frac{1}{h}),
$$

which imply that $T_6^* = o_p(h^{-1/2})$ and $T_7^* = o_p(h^{-1/2})$. Similarly, we can get $T_8^* = o_p(h^{-1/2})$ and $T_9^* = o_p(h^{-1/2})$. Note that T_2^*, T_4^* and T_5^* are the same as in Chapter 2 except replacing p by p_1, replacing X and $\Omega^*(U_t)$ respectively by $X^{(1)} - \Omega_{12}\Omega_{22}X^{(2)}$ and $\Omega_{11,2}$. Therefore, the remaining proof follows the same lines as those in the proof of Theorem 2.2.

Proof for Theorem 3.2: The proof of Theorem 3.2 is omitted since it is the special case of Theorem 3.1.

Proof for Theorem 3.3: The generalized quasi-likelihood ratio test statistic can be rewritten as

$$
\begin{aligned}
T_n &= \ell(H_a) - \ell(H_0) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_1(u)X_k^{(1)} - \hat{A}_2(u)X_k^{(2)} \right) \\
&\quad - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \widetilde{A}_1(u)X_k^{(1)} - \widetilde{A}_2(u)X_k^{(2)} \right) \\
&= \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_1(u)X_k^{(1)} - \hat{A}_2(u)X_k^{(2)} \right) \\
&\quad - \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_{10}(u)X_k^{(1)} - A_{20}(u)X_k^{(2)} \right) \\
&\quad + \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_{10}(u)X_k^{(1)} - A_{20}(u)X_k^{(2)} \right) \\
&\quad - \sum_{t=1}^{n} \rho_\tau \left(Y_t - \widetilde{A}_1(u)X_k^{(1)} - \widetilde{A}_2(u)X_k^{(2)} \right) \\
&= E_1 - E_2,
\end{aligned}
$$

where

$$
E_1 = \sum_{t=1}^{n} \rho_\tau \left(Y_t - \hat{A}_1(u)X_k^{(1)} - \hat{A}_2(u)X_k^{(2)} \right)
$$

$$-\sum_{t=1}^{n}\rho_\tau\left(Y_t - A_{10}(u)X_k^{(1)} - A_{20}(u)X_k^{(2)}\right)$$

and

$$E_2 = \sum_{t=1}^{n}\rho_\tau\left(Y_t - \tilde{A}_1(u)X_k^{(1)} - \tilde{A}_2(u)X_k^{(2)}\right)$$

$$-\sum_{t=1}^{n}\rho_\tau\left(Y_t - A_{10}(u)X_k^{(1)} - A_{20}(u)X_k^{(2)}\right).$$

For E_1, we can get the similar derivation as in Theorem 3.2 since $\hat{A}(u)$ is the non-parametric estimation using jackknife method and local linear fitting technique. Next, we focus on E_2.

$$E_2 = -\sum_{t=1}^{n}\begin{pmatrix}\tilde{A}_1 - A_{10}\\ \tilde{A}_2 - A_{20}\end{pmatrix}^T\begin{pmatrix}X_t^{(1)}\\ X_t^{(2)}\end{pmatrix}\varepsilon_t$$

$$+\frac{1}{2}\sum_{t=1}^{n}E(f_{y|u,x}(q_\tau|X_t,U_t))\begin{pmatrix}\tilde{A}_1 - A_{10}\\ \tilde{A}_2 - A_{20}\end{pmatrix}^T\begin{pmatrix}X_t^{(1)}\\ X_t^{(2)}\end{pmatrix}\begin{pmatrix}X_t^{(1)}\\ X_t^{(2)}\end{pmatrix}^T$$

$$\times\begin{pmatrix}\tilde{A}_1 - A_{10}\\ \tilde{A}_2 - A_{20}\end{pmatrix}$$

$$= -\sum_{t=1}^{n}(\tilde{A}_1 - A_{10})X_t^{(1)}\varepsilon_t - \sum_{t=1}^{n}(\tilde{A}_2 - A_{20})X_t^{(1)}\varepsilon_t + \frac{1}{2}\sum_{t=1}^{n}E(f_{y|u,x}(q_\tau|X_t,U_t))$$

$$[(\tilde{A}_1 - A_{10})^T X_t^{(1)}(X_t^{(1)})^T(\tilde{A}_1 - A_{10}) + (\tilde{A}_1 - A_{10})^T X_t^{(1)}(X_t^{(1)})^T(\tilde{A}_2 - A_{20})$$

$$(\tilde{A}_2 - A_{20})^T X_t^{(1)}(X_t^{(1)})^T(\tilde{A}_1 - A_{10}) + (\tilde{A}_2 - A_{20})^T X_t^{(1)}(X_t^{(1)})^T(\tilde{A}_2 - A_{20})].$$

From Lemmas 3.2 and 3.3, we know that

$$(\tilde{A}_1 - A_{10}) = O_p(1/\sqrt{n}) + O_p(h^2) \quad \text{and} \quad (\tilde{A}_2 - A_{20}) = O_p(1/\sqrt{nh}) + O_p(h^2).$$

Thus,

$$
\begin{aligned}
E_2 \;=\; & -\sum_{t=1}^{n}(\tilde{A}_2 - A_{20})X_t^{(1)}\varepsilon_t - O_p(\sqrt{h}) + \frac{1}{2}\sum_{t=1}^{n}E(f_{y|u,x}(q_\tau|X_t,U_t)) \\
& [(\tilde{A}_2 - A_{20})^T X_t^{(1)}(X_t^{(1)})^T(\tilde{A}_1 - A_{10}) + O_p(\sqrt{h}) + O_p(\sqrt{h}) + O_p(h)].
\end{aligned}
$$

Then, the rest of proof is similar to the proof of Theorem 3.2.

Proof for Theorem 3.4: Under the H_a, the true parameter $A_1(u) = A_{10}(U) + \frac{1}{\sqrt{nh}}\Delta(U)$. The test statistic T_n can be rewritten as

$$
\begin{aligned}
T_n \;=\; & \ell(H_a) - \ell(H_0) \\
=\; & \sum_{t=1}^{n}\rho_\tau\left(Y_t - \hat{A}_1(u)X_k^{(1)} - \hat{A}_2(u)X_k^{(2)}\right) \\
& -\sum_{t=1}^{n}\rho_\tau\left(Y_t - A_{10}X_k^{(1)} - \hat{A}_2(u)X_k^{(2)}\right) \\
=\; & \sum_{t=1}^{n}\rho_\tau\left(Y_t - \hat{A}_1(u)X_k^{(1)} - \hat{A}_2(u)X_k^{(2)}\right) \\
& -\sum_{t=1}^{n}\rho_\tau\left(Y_t - A_1 X_k^{(1)} - A_{20}(u)X_k^{(2)}\right) \\
& +\sum_{t=1}^{n}\rho_\tau\left(Y_t - A_1 X_k^{(1)} - A_{20}(u)X_k^{(2)}\right) \\
& -\sum_{t=1}^{n}\rho_\tau\left(Y_t - A_{10}X_k^{(1)} - A_{20}(u)X_k^{(2)}\right) \\
& +\sum_{t=1}^{n}\rho_\tau\left(Y_t - A_{10}X_k^{(1)} - A_{20}(u)X_k^{(2)}\right) \\
& -\sum_{t=1}^{n}\rho_\tau\left(Y_t - A_{10}X_k^{(1)} - \hat{A}_2(u)X_k^{(2)}\right) \\
\equiv\; & O_1 + O_2 + O_3.
\end{aligned}
$$

First, I consider O_2. By using Knight identity, we have

$$
\begin{aligned}
O_2 &= \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_1 X_k^{(1)} - A_{20}(u) X_k^{(2)} \right) \\
&\quad - \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_{10} X_k^{(1)} - A_{20}(u) X_k^{(2)} \right) \\
&= -[\sum_{t=1}^{n} \rho_\tau \left(Y_t - A_1 X_k^{(1)} - A_{20}(u) X_k^{(2)} \right) \\
&\quad - \sum_{t=1}^{n} \rho_\tau \left(Y_t - A_{10} X_k^{(1)} - A_{20}(u) X_k^{(2)} \right)] \\
&= -\frac{1}{\sqrt{nh}} \sum_{t=1}^{n} \Delta(U_t) X_t^{(1)} \varepsilon_t - \int_0^{-\frac{1}{\sqrt{nh}}\Delta(U_t)X_t^{(1)}} I_{\epsilon_t < s} - I_{\epsilon_t < 0} ds \\
&= -F_1 - F_2.
\end{aligned}
$$

By following the similar derivation in Theorem 2.2, we can get

$$
\begin{aligned}
F_2 &= \frac{1}{2nh} \sum_{t=1}^{n} E_z(f_{y|u,x}(q_\tau|X_t, U_t)) \Delta(U_t)^T X_t^{(1)} (X_t^{(1)})^T \Delta(U_t) \\
&= \frac{1}{2h} E(f_{y|u,x}(q_\tau|X_t, U_t) \Delta(U_t)^T X_t^{(1)} (X_t^{(1)})^T \Delta(U_t)) \\
&\equiv d_{2n}.
\end{aligned}
$$

For $O_1 + O_3$ we have the same result as theorem 3.2. Therefore,

$$
\begin{aligned}
T_n &= -\frac{1}{\sqrt{nh}} \sum_{t=1}^{n} \Delta(U_t) X_t^{(1)} \varepsilon_t - d_{2n} \\
&\quad - \sum_{t=1}^{n} \frac{1}{(\sqrt{nh})^2 f_u(U_t)} \sum_{i \neq t} \varepsilon_i (X_i^{(1)} - X_i^{(2)} \Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}(X_t^{(1)} \\
&\quad - X_t^{(2)} \Omega_{12}\Omega_{22}^{-1}) K(\frac{U_i - U_t}{h}) \varepsilon_t + \frac{1}{2} \sum_{t=1}^{n} E(f_{y|u,x}(q_\tau|X_t, U_t))(\frac{1}{nh f_u(u_0)})^2 \\
&\quad \sum_{t=1}^{n}\sum_{t=1}^{n} \varepsilon_i \varepsilon_j [(X_i^{(1)} - X_i^{(2)} \Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1} \\
&\quad (X_t^{(1)} - X_t^{(2)} \Omega_{12}\Omega_{22}^{-1})(X_t^{(1)} - X_t^{(2)} \Omega_{12}\Omega_{22}^{-1})(\Omega_{11,2})^{-1}
\end{aligned}
$$

$$(X_j^{(1)} - X_j^{(2)}\Omega_{12}\Omega_{22}^{-1})]K(\frac{U_t - u_0}{h})K(\frac{U_t - u_0}{h}) - T_2^* + T_4^* + T_5^* + T_6^* + T_7^*$$
$$+ T_8^* + T_9^*,$$

where T_2^*, T_4^*, T_5^*, T_6^*, T_7^*, T_8^* and T_9^* are defined in the proof of Theorem 3.2 which are dominated by d_{1n}. So, The rest of proof is similar to the proof of Theorem 2.8 and Theorem 3.2. The details are omitted. $\square$

CHAPTER 4: CONCLUSION

In this dissertation, I propose some new test procedures, termed as generalized quasi-likelihood ratio test, to testing some hypotheses for varying coefficient quantile regression models, such as testing whether coefficients or partial coefficients are indeed varying or of some specific functional form.

First, I use local linear technique with jackknife method to estimate the non-parametric coefficient functions and derive the Bahadur representation of the estimators. Then, I propose the new test statistics which are constructed based on the comparison of the quasi-likelihood under null and alternative hypotheses to test hypotheses about the form of coefficients for varying coefficient quantile regression models. I also conduct some Monte Carlo simulation studies to illustrate the power of the proposed test procedure and an application to a real data set is also reported.

Also, I apply this test procedure to test the hypothesis on partial coefficients for varying coefficient quantile regression models. I adopt the similar test statistic to test whether partial coefficients are constant or of some specific functional form with other coefficients remaining completely unspecified. The asymptotic distributions of the proposed test statistics are derived and some simulation results are presented to show the effectiveness of the proposed test procedure.

There are still some interesting research topics related to this dissertation which deserve further investigation. First, one may release the constriction of i.i.d assumption. In my dissertation, I only derive the asymptotic result under the i.i.d setting. It should hold for the non i.i.d data, such as stationary time series data or under some mixing conditions. Secondly, the generalized quasi-likelihood ratio test statis-

tic can be extended to the hypothesis problem for other models, such as additive quantile model, predictive regression model and others. Furthermore, there is few paper available in the literature about varying coefficient quantile regression models under the nonstationary time series setting due to the difficulty of deriving Bahadur representation for the nonstationary time series data. All of the aforementioned issues can be considered as future research topics.

REFERENCES

[1] Bassett, G. and Chen, H. (2001). Portfolio Style: Return-Based Attribution Using Quantile Regression Empirical Economics, Springer-Verlag, 1405-1441.

[2] Bassett, G. W. and Koenker, R. W. (1982). An Empirical Quantile Function for Linear Models with iid Errors. Journal of the American Statistical Association, 77, 407-415.

[3] Barnes, M. and Hughes, A. (2002). A Quantile Regression Analysis of the Cross Section of Stock Market Returns, Working paper, FederalReserve Bank of Boston.

[4] Cai, Z. (2002). Regression Quantile for Time Series. Econometric Theory, 18, 169-192.

[5] Cai, Z. (2010). Functional Coefficient Models for Economic and Financial data. In Ferraty, F., Romain, Y.(Eds), *Handbook of Functional Data Analysis*, Oxford University Press, Oxford, 166-186.

[6] Cai, Z., Das, M., Xiong, H. and Wu, X.(2006). Functional Coefficient Instrumental Variables Models. Journal of Econometrics, 133, 207-241.

[7] Cai, Z., Gu, J. and Li, Q. (2009). Some Recent Development on Nonparametric Econometrics. Advances in Econometrics, 25, 495-549.

[8] Cai, Z., Fan, J. and Li, R. (2000). Efficient Estimation and Inferences for Varying-Coefficient Models. Journal of the American Statistical Association, 95, 888-902.

[9] Cai, Z., Fan, J. and Yao, Q. (2000). Functional-Coefficient Regression Models for Nonlinear Time Series. Journal of the American Statistical Association, 95, 941-956.

[10] Cai, Z. and Xu, X. (2008). Nonparametric Quantile Estimations for Dynamic Smooth Coefficient Models. Journal of the American Statistical Association, 103, 1595-1608.

[11] Cai, Z. and Xiao, Z.(2012). Semiparametric Quantile Regression Estimation in Dynamic Models with Partially Varying Coefficients. Journal of Econometrics, 167, 413-425.

[12] Chaudhuri, P. (1991). Nonparametric Quantile Estimates of Regression Quantiles and Their Local Bahadur Representation. The Annals of Statistics, 19, 760-777.

[13] Chen, X. and Linton, O. (2001). The Estimation of Conditional Densities. LSE STICERD Research Paper NO. EM415.

[14] Cleveland, W.S., Grosse, E. and Shyu, W.M. (1991). Local Regression Models. In Statistical Models in S (Chambers, J. M. and Hastie, T. J., eds), 309-76. Wadsworth & Brooks, Pacific Grove.

[15] Duffie, D. and Pan,J. (1997). An Overview of Value at Risk. Journal of Derivatives, 4, 7-49.

[16] De Jong, P. (1987). A Central Limit Theorem for Generalized Quadratic Forms. Probability Theory and Related Fields, 75, 261-277.

[17] Efron, B. (1991). Regression Percentiles Using Asymmetric Squared Error Loss. Statistica Sinica, 1, 93-125.

[18] El Bantli, F. and Hallin, M. (1999). L_1 Estimation in Linear Models with Heterogeneous White Noise. Statistics and Probability Letters, 45, 305-315.

[19] Engle, R. and Manganelli, S. (2004). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. Journal of Business and Economic Statistics, 22, 367-381.

[20] Fan, J. and Gijbels, I. (1996). *Local Polynomial Modeling and Its Applications*. Chapman and Hall, London.

[21] Fan, J., Yao, Q. and Tong. H. (1996). Estimation of Conditional Densities and Sensitivity Measures in Nonlinear Dynamical Systems. Biometrika, 83, 189-206.

[22] Fan, J. and Zhang, W. (2008). Statistical Methods with Varying Coefficient Models. Statistic Interface, 1, 179-195.

[23] Fan, J., Zhang, C. and Zhang, J. (2001). Generalized Likelihood Ratio Statistics and Wilks Phenomenon. The Annuals of Statistics, 29, 153-193.

[24] illey, O. W.and Pace, R. K.(1996). On the Harrison and Rubinfeld Data. Journal of Environmental Economics and Management, 31, 403-405.

[25] Gourieroux, C., Monfort, A. and Renault, E. (1987). Consistent M-Estimators in a Semi-Parametric Model. Document de Travail de INSEE, 8706.

[26] Granger, C.W.J., White, H. and Kamstra, M. (1989). Interval Forecasting: an Analysis Based upon ARCH-quantile Estimates. Journal of Econometrics, 40, 87-96.

[27] Hastie, T. and Tibshirani, R. (1993). Varying-coefficient Models. Journal of Royal Statistical Society Series B, 55, 757-796.

[28] Harrison, D. and Rubinfeld, D.L. (1978). Hedonic Prices and the Demand for Clean Air. Journal of Environmental Economics and Management, 5, 81-102.

[29] He, X. and Shi, P. (1996). Bivariate Tensor-product B-splines in a Partly Linear Model. Journal of Multivariate Analysis, 58, 162-181.

[30] He, X., and Ng, P. (1999). Quantile Splines with Several Covariates. Journal of Statistical Planning and Inference, 75, 343-352.

[31] He,X., Ng, P. and Portnoy, S. (1998). Bivariate Quantile Smoothing Splines. Journal of the Royal Statistical Society Series B, 60, 537-550.

[32] He, X. and Zhu, L.-X. (2003). A Lack-of-fit Test for Quantile Regression. Journal of the American Statistical Association, 98, 1013-1022.

[33] Hendricks, W. and Koenker, R. (1992). Hierarchical Spline Models for Conditional Quantiles and the Demand for Electricity. Journal of the American Statistical Association, 87, 58-68.

[34] Honda, T. (2004). Quantlie Regression in Varying Coedfficient Models. Journal of Statistical Planning and Inference, 121, 113-125.

[35] Horowitz, J.L. and Spokoiny, V.G. (2002). An Adaptive Rate-optimal Test of Linearity for Median Regression Models. Journal of the American Statistical Association, 97, 822-835.

[36] Horowitz, J.L. and Lee, S. (2005). Nonparametric Estimation of an Additive Quantile Regression Model. Journal of the American Statistical Association, 100, 1238-1249.

[37] Huber, P. J. (1964). Robust Estimation of a Location Parameter. Annals of Mathematical Statistics, 35, 73-101.

[38] Khindanova, I. N. and Rachev, S. T. (2000). Value at Risk: Recent Advances. *Handbook on Analytic-Computational Mehtods in Applied Mathematics*. CRC Press LLC.

[39] Koenker R. (2000). Galton, Edgeworth, Frisch, and Prospects for Quantile Regression In Econometrics. Journal of Econometrics, 95, 347-374.

[40] Koenker, R. (2004). Quantreg: An R Package for Quantile Regression and Related Methods. http://cran.r-project.org.

[41] Koenker, R. (2005). *Quantile Regression*. Econometric Society Monograph Series, Cambridge University Press, New York.

[42] Koenker, R. and Bassett, G. (1978). Regression Quantiles. Econometrica, 46, 33-50.

[43] Koenker, R. and Bassett, G. (1982a). Robust Tests for Heteroscedasticity Based on Regression Quantiles. Econometrica, 50, 43-61.

[44] Koenker, R. and Bassett, G. (1982b). Test of Linear Hypotheses and L_1 Estimation. Econometrica, 50, 1577-1584.

[45] Koenker, R. and Hallock, K.F.(2001). Quantile Regression: An Introduction. Journal of Economic Perspectives, 15, 143-157.

[46] Koenker, R. and Machado, J. (1999). Goodness of Fit and Related Inference Processes for Quantile Regression. Journal of the Royal Statistical Society Series B, 66, 145-163.

[47] Koenker, R. and Xiao, Z. (2002). Inference on the Quantile Regression Process. Econometrica, 70, 1583-1612.

[48] Koenker, R. and Xiao, Z. (2004). Unit Root Quantile Autoregression Inference. Journal of the American Statistical Association, 99, 775-787.

[49] Koenker, R. and Xiao, Z. (2006). Quantile Autoregression, with Discussion and Rejoinder. Journal of the American Statistical Association, 96, 980-1006.

[50] Koenker, R. and Zhao, Q. (1996). Conditional Quantile Estimation and Inference for ARCH Models. Econometric Theory, 12, 793-813.

[51] Komunjer, I. (2005). Quasi-maximum Likelihood Estimation for Conditional Quantiles. Journal of Econometrics, 128, 137-164.

[52] Morgan, J. P. (1995). *Riskmetrics Technical Manual*, 3rd Edition.

[53] Powell, J. L. (1986). Censored Regression Quantiles. Journal of Econometrics, 32, 143-155.

[54] Roncheti, E. (1985). Robust Model Selection in Regression. Statistics &Probability Letters, 3, 21-23.

[55] Rose, S.J. (1992). *Social Stratification in the United States*. The New Press, New York.

[56] Taylor, J.W. and Bunn, D. W. (1999). A quantile regression approach to generating prediction intervals. Management Science, 45, 225-237.

[57] Tsay, R.S. (2000). Extreme Values, Quantile Estimation and Value at Risk. Working paper, Graduate School of Business, University of Chicago.

[58] Tsay, R.S. (2002). *Analysis of Financial Time Serise*. 1th Edition, John Wiley & Sons, New York.

[59] Wu, T., Yu, K. and Yu, Y. (2010). Single-index Quantile Regression. Journal of Multivariate Analysis, 101, 1607-1621.

[60] Xu, X. (2005). *Semi/Non-parametric Dynamic Quantile Regression Models*. Ph.D. Dissertation. University of North Carolina at Charlotte.

[61] Yu, K. and Lu, Z. (2004). Local Linear Additive Quantile Regression. Scandinavian Journal of Statistics, 31, 333-346.

[62] Zhang, C. (2000). Topics in Generalized Likelihood Ratio Test. Ph.D. Dissertation, The University of North Carolina at Chapel Hill.