Econometric Theory, **28**, 2012, 861–887.
doi:10.1017/S0266466611000685

A CONSISTENT NONPARAMETRIC TEST FOR CAUSALITY IN QUANTILE

Kiho Jeong
Kyungpook National University

Wolfgang K. Härdle
Humboldt-Universität zu Berlin

Song Song
Humboldt-Universität zu Berlin
and
University of California, Berkeley

This paper proposes a nonparametric test of Granger causality in quantile. Zheng (1998, *Econometric Theory* 14, 123–138) studied the idea to reduce the problem of testing a quantile restriction to a problem of testing a particular type of mean restriction in independent data. We extend Zheng's approach to the case of dependent data, particularly to the test of Granger causality in quantile. Combining the results of Zheng (1998) and Fan and Li (1999, *Journal of Nonparametric Statistics* 10, 245–271), we establish the asymptotic normal distribution of the test statistic under a β-mixing process. The test is consistent against all fixed alternatives and detects local alternatives approaching the null at proper rates. Simulations are carried out to illustrate the behavior of the test under the null and also the power of the test under plausible alternatives. An economic application considers the causal relations between the crude oil price, the USD/GBP exchange rate, and the gold price in the gold market.

1. INTRODUCTION

Whether movements in one economic variable cause reactions in another variable is an important issue in economic policy and also for financial investment decisions. A framework for investigating causality between economic indicators has been developed by Granger (1969). Testing for Granger causality between

The research was conducted while Jeong was visiting C.A.S.E.—Center for Applied Statistics and Economics—Humboldt-Universität zu Berlin in the summers of 2005 and 2007. Jeong is grateful for their hospitality during the visits. Jeong's work was supported by a Korean Research Foundation grant funded by the Korean government (MOEHRD) (KRF-2006-B00002), and Härdle and Song's work was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk." We thank the editor, two anonymous referees, and Holger Dette for concrete suggestions on improving the manuscript and restructuring the paper. Their valuable comments and suggestions are gratefully acknowledged. Address correspondence to Kiho Jeong, Kyungpook National University, Korea; e-mail: khjeong@knu.ac.kr.

economic time series has been since studied intensively in empirical macroeconomics and empirical finance. The majority of research results were obtained in the context of Granger causality in the conditional mean. The conditional mean, though, is a questionable element of analysis if the distributions of the variables involved are nonelliptic or fat tailed as is to be expected with, for example, financial returns. The focus of a causality analysis on the mean might result in unclear news. The conditional mean is only one element of an overall summary for the conditional distribution. A tail area causal relation may be quite different from a causality based on the center of the distribution. Lee and Yang (2007) explore money-income Granger causality in the conditional quantile and find that Granger causality is significant in tail quantiles, whereas it is not significant in the center of the distribution.

An illustrating motivation for the research presented here is from labor market analysis where one tries to find out how income depends on the age of the employee for different education levels, genders, and nationalities, and so on (discrimination effects); see, for example, Buchinsky (1995). In particular, the effect of education on income is summarized by the basic claim of Day and Newburger (2007): At most ages, more education equates with higher earnings, and the payoff is most notable at the highest educational level, which is actually from the point of view of mean regression. However, whether this difference is significant or not is still questionable, especially for different ends of the (conditional) income distribution. Härdle, Ritov, and Song (2009) show that for the 0.20 quantile confidence bands for income given "university," "apprenticeship," and "low education" status do not differ significantly from one another although they become progressively lower, which indicates that high education does not equate to higher earnings significantly for the lower tails of income, whereas increasing age seems to be the main driving force. For the conditional median, the bands for "university" and "low education" differ significantly. For the 0.80 quantiles, all conditional quantiles differ, which indicates that higher education is associated with higher earnings. However, these findings do not necessarily indicate causalities. To answer the question "Does education Granger cause income in various conditional quantiles?" the concept of Granger causality in means cannot be used to estimate or test for these effects. Hence the need for the concept of Granger causality in quantiles and the need to develop tests for these effects emerge.

Another motivation comes from controlling and monitoring downside market risk and investigating large comovements between financial markets. These are important for risk management and portfolio/investment diversification (Hong, Liu, and Wang, 2009). Various other risk management tasks are described in Bollerslev (2001) and Campbell and Cochrane (1999) indicating the importance of Granger causality in quantile. Yet another motivation comes from the well-known robustness properties of the conditional quantile: like the parallel boxplot—calculated across an explanatory variable—the set of conditional quantiles characterizes the entire distribution in more detail.

Based on the kernel method, we propose a nonparametric test for Granger causality in quantile. Testing conditional quantile restrictions by nonparametric estimation techniques in dependent data situations has not been considered in the literature before. This paper intends to fill this literature gap. In an unpublished working paper that has been independently carried out from ours, Lee and Yang (2007) also propose a test for Granger causality in the conditional quantile. Their test, however, relies on linear quantile regression and thus is subject to possible functional misspecification of quantile regression. Recently, Hong et al. (2009) investigated Granger causality in value at risk (VaR) with a corresponding (kernel-based) test. Their method, however, offers two possible improvements. The first is that it needs a parametric specification of VaR, again subject to misspecification errors. The second is that their test does not directly check causality but rather a necessary condition for causality.

The problem of testing conditional mean restrictions using nonparametric estimation techniques has been actively studied for dependent data. Among the related work, the testing procedures of Fan and Li (1999) and Li (1999) use the general hypothesis of the form $\mathrm{E}(\varepsilon|z) = 0$, where ε and z are the regression error term and the vector of regressors, respectively. They consider the distance measure of $J = \mathrm{E}[\varepsilon \mathrm{E}(\varepsilon|z) f(z)]$ to construct kernel-based procedures. For the advantages of using this distance measure in kernel-based testing procedures, see Li and Wang (1998) and Hsiao and Li (2001). A feasible test statistic based on J has a second-order degenerate U-statistic as the leading term under the null hypothesis. Generalizing the result of Hall (1984) for independent data, Fan and Li (1999) establish the asymptotic normal distribution for a general second-order degenerate U-statistic with dependent data.

All the results stated previously on testing mean restrictions are however irrelevant when testing quantile restrictions. Zheng (1998) proposed an idea to transform quantile restrictions to mean restrictions in independent data. Following his idea, one can use the existing technical results on testing mean restrictions in testing quantile restrictions. In this paper, by combining Zheng's idea and the results of Fan and Li (1999) and Li (1999), we derive a test statistic for Granger causality in quantile and establish the asymptotic normal distribution of the proposed test statistic under a β-mixing process. Our testing procedure can be extended to several hypothesis testing problems with conditional quantile in dependent data; for example, testing a parametric regression functional form, testing the insignificance of a subset of regressors, and testing semiparametric versus nonparametric regression models.

The paper is organized as follows. Section 2 presents the test statistic. Section 3 establishes the asymptotic normal distribution under the null hypothesis of no causality in quantile. Section 4 displays a fairly extensive simulation study to illustrate the behavior of the test under the null, in addition to the power of the test under plausible alternatives. Section 5 considers the causal relations between the crude oil and gold prices as an economic application. Section 6 concludes the paper. Technical proofs are given in the Appendix.

2. NONPARAMETRIC TEST FOR GRANGER CAUSALITY IN QUANTILE

To simplify the exposition, we assume a bivariate case, or that only $\{y_t, w_t\}$ are observable. Granger causality in mean (Granger, 1988) is defined as follows.

1. w_t does not cause y_t in mean with respect to $\{y_{t-1}, \ldots, y_{t-p}, w_{t-1}, \ldots, w_{t-q}\}$ if

$$\mathrm{E}(y_t | y_{t-1}, \ldots, y_{t-p}, w_{t-1}, \ldots, w_{t-q}) = \mathrm{E}(y_t | y_{t-1}, \ldots, y_{t-p}) \quad \text{and}$$

2. w_t is a prima facie cause in mean of y_t with respect to $\{y_{t-1}, \ldots, y_{t-p}, w_{t-1}, \ldots, w_{t-q}\}$ if

$$\mathrm{E}(y_t | y_{t-1}, \ldots, y_{t-p}, w_{t-1}, \ldots, w_{t-q}) \neq \mathrm{E}(y_t | y_{t-1}, \ldots, y_{t-p}).$$

Motivated by the definition of Granger causality in mean, we define Granger causality in quantile as follows.

1. w_t does not cause y_t in the θ-quantile with respect to $\{y_{t-1}, \ldots, y_{t-p}, w_{t-1}, \ldots, w_{t-q}\}$ if

$$Q_\theta(y_t | y_{t-1}, \ldots, y_{t-p}, w_{t-1}, \ldots, w_{t-q}) = Q_\theta(y_t | y_{t-1}, \ldots, y_{t-p}). \tag{1}$$

2. w_t is a prima facie cause in the θ-quantile of y_t with respect to $\{y_{t-1}, \ldots, y_{t-p}, w_{t-1}, \ldots, w_{t-q}\}$ if

$$Q_\theta(y_t | y_{t-1}, \ldots, y_{t-p}, w_{t-1}, \ldots, w_{t-q}) \neq Q_\theta(y_t | y_{t-1}, \ldots, y_{t-p}), \tag{2}$$

where $Q_\theta(y_t|\cdot)$ is the θth ($0 < \theta < 1$) conditional quantile of y_t given $\cdot$, which depends on t.

Denote $x_t \equiv (y_{t-1}, \ldots, y_{t-p})$, $z_t \equiv (y_{t-1}, \ldots, y_{t-p}, w_{t-1}, \ldots, w_{t-q})$, and the conditional distribution function y_t given $z_t(x_t)$ by $F_{y_t|z_t}(y_t|z_t)(F_{y_t|z_t}(y_t|x_t))$, which is abbreviated as $F_{y|z}(y|z)$ $(F_{y|x}(y|x))$ later, and $v_t = (x_t, z_t)$. In this paper, $F_{y|z}(y|z)$ is assumed to be absolutely continuous in y for almost all $v = (x, z)$. Denote $Q_\theta(z_t) \equiv Q_\theta(y_t|z_t)$ and $Q_\theta(x_t) \equiv Q_\theta(y_t|x_t)$. Then we have, with probability 1,

$$F_{y|z}\{Q_\theta(z_t)|z_t\} = \theta, \qquad v = (x, z) \quad \text{and}$$

from the definitions (1) and (2), the hypotheses to be tested are

$$\mathrm{H}_0 : \mathrm{P}\left\{F_{y|z}(Q_\theta(x_t)|z_t) = \theta\right\} = 1 \quad \text{a.s.} \tag{3}$$

$$\mathrm{H}_1 : \mathrm{P}\left\{F_{y|z}(Q_\theta(x_t)|z_t) = \theta\right\} < 1 \quad \text{a.s.} \tag{4}$$

Zheng (1998) proposed an idea to reduce the problem of testing a quantile restriction to a problem of testing a particular type of mean restriction. The null hypothesis (3) is true if and only if $\mathrm{E}[1\{y_t \leqslant Q_\theta(x_t)|z_t\}] = \theta$ or $1\{y_t \leqslant Q_\theta(x_t)\} = \theta + \varepsilon_t$ where $\mathrm{E}(\varepsilon_t|z_t) = 0$ and $1(\cdot)$ is the indicator function. For a list of related literature we refer to Li and Wang (1998) and Zheng (1998). Although various distance measures can be used to consistently test the hypothesis (3), we consider the following distance measure:

$$J \equiv \mathrm{E}\left[\left\{F_{y|z}(Q_\theta(x_t)|z_t) - \theta\right\}^2 f_z(z_t)\right], \tag{5}$$

with $f_{z_t}(z_t)$ being the marginal density function of z_t, which is sometimes abbreviated as $f_z(z_t)$. Note that $J \geqslant 0$ and the equality holds if, and only if, H_0 is true, with strict inequality holding under H_1. Thus J can be used as a proper candidate for consistent testing of H_0 (Li, 1999, p. 104). Because $\mathrm{E}(\varepsilon_t|z_t) = F_{y|z}\{Q_\theta(x_t)|z_t\} - \theta$ we have

$$J = \mathrm{E}\{\varepsilon_t \mathrm{E}(\varepsilon_t|z_t) f_z(z_t)\}. \tag{6}$$

The test is based on a sample analogue of $\mathrm{E}\{\varepsilon\, \mathrm{E}(\varepsilon|z) f_z(z)\}$. Including the density function $f_z(z)$ avoids the problem of trimming on the boundary of the density function; see Powell, Stock, and Stoker (1989) for an analogue approach. The density-weighted conditional expectation $\mathrm{E}(\varepsilon|z) f_z(z)$ can be estimated by kernel methods

$$\hat{\mathrm{E}}(\varepsilon_t|z_t)\hat{f}_z(z_t) = \frac{1}{(T-1)h^m} \sum_{s \neq t}^{T} K_{ts}\varepsilon_s, \tag{7}$$

where $m = p + q$ is the dimension of z, $K_{ts} = K\{(z_t - z_s)/h\}$ is the kernel function, and h is a bandwidth. Then we have a sample analogue of J as

$$\begin{aligned} J_T &\equiv \frac{1}{T(T-1)h^m} \sum_{t=1}^{T} \sum_{s \neq t}^{T} K_{ts}\varepsilon_t\varepsilon_s \\ &= \frac{1}{T(T-1)h^m} \sum_{t=1}^{T} \sum_{s \neq t}^{T} K_{ts}[1\{y_t \leqslant Q_\theta(x_t)\} - \theta][1\{y_s \leqslant Q_\theta(x_s)\} - \theta]. \end{aligned} \tag{8}$$

The θth conditional quantile of y_t given x_t, $Q_\theta(x_t)$, can also be estimated by the nonparametric kernel method

$$\hat{Q}_\theta(x_t) = \hat{F}_{y|x}^{-1}(\theta|x_t), \tag{9}$$

where

$$\hat{F}_{y|x}(y_t|x_t) = \frac{\sum_{s \neq t} L_{ts} 1(y_s \leqslant y_t)}{\sum_{s \neq t} L_{ts}} \tag{10}$$

is the Nadaraya–Watson kernel estimator of $F_{y|x}(y_t|x_t)$ with the kernel function of $L_{ts} = L(x_t - x_s)/a$ and the bandwidth parameter of α. The unknown error ε can be estimated as

$$\hat{\varepsilon}_t \equiv I\left\{y_t \leqslant \hat{Q}_\theta(x_t)\right\} - \theta. \tag{11}$$

Replacing ε by $\hat{\varepsilon}$, we have a feasible kernel-based test statistic of J,

$$\begin{aligned}\hat{J}_T &\equiv \frac{1}{T(T-1)h^m}\sum_{t=1}^{T}\sum_{s\neq t}^{T} K_{ts}\hat{\varepsilon}_t\hat{\varepsilon}_s \\ &= \frac{1}{T(T-1)h^m}\sum_{t=1}^{T}\sum_{s\neq t}^{T} K_{ts}\left[1\left\{y_t \leqslant \hat{Q}_\theta(x_t)\right\} - \theta\right]\left[1\left\{y_s \leqslant \hat{Q}_\theta(x_s)\right\} - \theta\right].\end{aligned} \tag{12}$$

3. THE LIMITING DISTRIBUTIONS OF THE TEST STATISTIC

Two existing works are useful in deriving the limiting distribution of the test statistic; one is Theorem 2.3 of Franke and Mwita (2003) on the uniform convergence rate of a nonparametric quantile estimator; another is Lemma 2.1 of Li (1999) on the asymptotic distribution of a second-order degenerate U-statistic, which is derived from Theorem 2.1 of Fan and Li (1999). We restate these results in lemmas subsequently for ease of reference. We collect the assumptions needed for Theorem 3.1.

(A1)

(a) $\{y_t, w_t\}_{t=1}^T$ is strictly stationary and absolutely regular with mixing coefficients $\beta(\tau) = \mathcal{O}(\rho^\tau)$ for some $0 < \rho < 1$.

(b) For some integer $v \geqslant 2$, f_y, f_z, and f_x all are bounded and belong to $\mathfrak{A}_v^\infty$ (see (D2) later in this section).

(c) Use $\mu_s^t(z)$ $(\mu_s^t(\varepsilon))$ to denote the σ algebra generated by $(z_s, ..., z_t)$ $((\varepsilon_s, ..., \varepsilon_t))$ for $s \leq t$. With probability 1, $\mathrm{E}\left[\varepsilon_t|\mu_{-\infty}^t(z), \mu_{-\infty}^{t-1}(\varepsilon)\right] = 0$, that is, the error ε_t is a martingale difference process. The terms $\mathrm{E}\left[\left|\varepsilon_t^{4+\eta}\right|\right] < \infty$ and $\mathrm{E}\left[\left|\varepsilon_{t_1}^{i_1}\varepsilon_{t_2}^{i_2}\ldots\varepsilon_{t_l}^{i_l}\right|^{1+\xi}\right] < \infty$ for some arbitrarily small $\eta > 0$ and $\xi > 0$, where $2 \leqslant l \leqslant 4$ is an integer, $0 \leqslant i_j \leqslant 4$, and $\sum_{j=1}^l i_j \leqslant 8$. The terms $\sigma_\varepsilon^2(z) = \mathrm{E}(\varepsilon_t^2|z_t = z)$ and $\mu_{\varepsilon 4}(z) = \mathrm{E}\left[\varepsilon_t^4|z_t = z\right]$ all satisfy some Lipschitz conditions: $|a(u+v) - a(u)| \leqslant D(u)\|v\|$ with $\mathrm{E}\left[|D(z)|^{2+\eta'}\right] < \infty$ for some small $\eta' > 0$, where $a(\cdot) = \sigma_\varepsilon^2(\cdot), \mu_{\varepsilon 4}(\cdot)$.

(d) Let $f_{\tau_1,...,\tau_l}(\,)$ be the joint probability density function of $\left(z_{\tau_1}, \ldots, z_{\tau_l}\right)$ $(1 \leqslant l \leqslant 3)$. Then $f_{\tau_1,...,\tau_l}(\,)$ is bounded and satisfies a Lipschitz condition: $\left|f_{\tau_1,...,\tau_l}(z_1+u_1, z_2+u_2, \ldots, z_l+u_l) - f_{\tau_1,...,\tau_l}(z_1, z_2, \ldots, z_l)\right| \leqslant D_{\tau_1,...,\tau_l}(z_1, \ldots, z_l)\|u\|$, where $u = (u_1, ..., u_l)$, $z = (z_1, ..., z_l)$, and $D_{\tau_1,...,\tau_l}(\,)$ is integrable and satisfies the condition that $\int\int\int D_{\tau_1,...,\tau_l}(z_1, \ldots, z_l)\|z\|^{2\xi} dz_1, \ldots, dz_l < M < \infty$ and $\int\int\int D_{\tau_1,...,\tau_l}(z_1, \ldots, z_l) f_{\tau_1,...,\tau_l}(z_1, \ldots, z_l) dz_1, ..., dz_l < M < \infty$ for some $\xi > 1$.

(e) For any y and x satisfying $0 < F_{y|x}(y|x) < 1$ and $f_x(x) > 0, F_{y|x}$ and $f_x(x)$ are continuous and bounded in x and y; for fixed y, the conditional distribution function $F_{y|x}$ and the conditional density function $f_{y|x}$ belong to $\mathfrak{A}_3^\infty$; $f_{y|x}(Q_\theta(x)|x) > 0$ for all x; $f_{y|x}$ is uniformly bounded in x and y by, say, c_f.

(f) For some compact set G, there are $\varepsilon > 0$ and $\gamma > 0$ such that $f_x \geqslant \gamma$ for all x in the ε-neighborhood $\{x| \|x-u\| < \varepsilon,\ u \in G\}$ of G. For the compact set G and some compact neighborhood Θ_0 of 0, the set $\Theta = \{v = Q_\theta(x) + \mu | x \in G, \mu \in \Theta_0\}$ is compact, and for some constant $c_0 > 0$, $f_{y|x}(y|x) \geqslant c_0$ for all $x \in G, v \in \Theta$.

(g) There is an increasing sequence s_T of positive integers such that for some finite A,

$\frac{T}{s_T}\beta^{2s_T/(3T)}(s_T) \leqslant A, \qquad 1 \leqslant s_T \leqslant \frac{T}{2} \qquad \text{for all } T \geqslant 1.$

(A2)

(a) We use product kernels for both $L(\cdot)$ and $K(\cdot)$. Let l and k be their corresponding univariate kernel which is bounded and symmetric. Then $l(\cdot)$ is nonnegative, $l(\cdot) \in \Upsilon_v$, $k(\cdot)$ is nonnegative, and $k(\cdot) \in \Upsilon_2$.

(b) $h = \mathcal{O}(T^{-\alpha'})$ for some $0 < \alpha' < (7/8)m$.

(c) $a = \mathcal{O}(1)$ and $\tilde{S}_T = Ta^p(s_T \log T)^{-1} \to \infty$ for some $s_T \to \infty$.

(d) A positive number δ exists such that for $r = 2+\delta$ and a generic number M_0

$$\int\int \left|\frac{1}{h^m}K\left(\frac{z_1 - z_2}{h}\right)\right|^r dF_z(z_1)dF_z(z_2) \leqslant M_0 < \infty \quad \text{and}$$

$$\mathrm{E}\left|\frac{1}{h^m}K\left(\frac{z_1 - z_2}{h}\right)\right|^r \leqslant M_0 < \infty.$$

(e) For some δ' $(0 < \delta' < \delta)$, $\beta(T) = \mathcal{O}(T^{-(2+\delta')/\delta'})$.

The following definitions are due to Robinson (1988).

DEFINITION (D1). Υ_λ, $\lambda \geqslant 1$ *is the class of even functions* $k: R \to R$ *satisfying* $\int_R u^i k(u)du = \delta_{i0}$ $(i = 0, 1, \ldots, \lambda - 1)$,

$$k(u) = \mathcal{O}\left((1+|u|^{\lambda+1+\varepsilon})^{-1}\right), \quad \textit{for some } \varepsilon > 0,$$

where δ_{ij} *is the Kronecker's delta.*

DEFINITION (D2). $\mathfrak{A}_\mu^\alpha$, $\alpha > 0$, $\mu > 0$ *is the class of functions* $g: R^m \to R$ *satisfying that* g *is* $(d-1)$*-times partially differentiable for* $d - 1 \leqslant \mu \leqslant d$; *for some* $\rho > 0$, $\sup_{y \in \phi_{z\rho}} \left|g(y) - g(z) - G_g(y,z)\right| / |y-z|^\mu \leqslant D_g(z)$ *for all* z, *where* $\phi_{z\rho} = \{y| |y-z| < \rho\}$; $G_g = 0$ *when* $d = 1$; G_g *is a* $(d-1)th$ *degree*

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

homogeneous polynomial in $y-z$ *with coefficients being the partial derivatives of* g *at* z *of orders 1 through* $d-1$ *when* $d>1$; *and* $g(z)$, *its partial derivatives of order* $d-1$ *and less, and* $D_g(z)$ *have finite* αth *moments.*

The functions in $\mathfrak{A}_\mu^\alpha$ are thus expanded in a Taylor series with a local Lipschitz condition on the remainder, (α, μ) depending simultaneously on smoothness and moment properties. Bounded functions in Lip(μ) (the Lipschitz class of degree μ) for $0<\mu\leq 1$ are in $\mathfrak{A}_\mu^\alpha$; for $\mu>1$, $\mathfrak{A}_\mu^\alpha$ contains the bounded and $(d-1)$-times boundedly differentiable functions whose $(d-1)$th partial derivatives are in Lip $(\mu-d+1)$). In applying $\mathfrak{A}_\mu^\alpha$ to f and F, we take $\alpha=\infty$.

Conditions (A1)(a)–(d) and (A2)(a) and (b) are adopted from conditions (D1) and (D2) of Li (1999), which are used to derive the asymptotic normal distribution of a second-order degenerate U-statistic. Assumption (A1)(a) requires $\{y_t, w_t\}_{t=1}^T$ to be a stationary absolutely regular process with geometric decay rate. Assumptions (A1)(b)–(d) are mainly some smoothness and moment conditions; these conditions are quite weak in the sense that they are similar to those used in Fan and Li (1996) for the independent data case. However, for autoregressive conditionally skedastic (ARCH) or generalized autoregressive conditionally heteroskedastic (GARCH) type error processes as considered in Engle (1982) and Bollerslev (1986), the error term ε_t may not have finite fourth moments in some situations. For example, let $\varepsilon_t|\varepsilon_{t-1}\sim N(0,\alpha_0+\alpha_1\varepsilon_{t-1}^2)$. Engle (1982) showed that ε_t does not have a finite fourth moment if $\alpha_1>1/\sqrt{3}$. Thus, Assumption (A1)(c) will be violated in such a case.

Assumption (A2)(a) requires $L(\cdot)$ to be a υth- ($\upsilon\geqslant 2$) order kernel. This condition together with (A1)(b) ensures that the bias in the kernel estimation (of the null model) is bounded. The requirement that k is a nonnegative second-order kernel function in (A2)(b) is a quite weak and standard assumption.

Conditions (A1)(e)–(g) and (A2)(c) are technical conditions (A1), (A2), (B1), (B2), (C1), and (C2) of Theorem 2.3 of Franke and Mwita (2003), which are required to get the uniform convergence rate of the nonparametric kernel estimator of the conditional distribution function and corresponding conditional quantile with mixing data. Because the simple ARCH models (Engle, 1982; Masry and Tjøstheim, 1995, 1997), their extensions (Diebolt and Guegan, 1993), and the bilinear Markovian models are geometrically strongly mixing under some general ergodicity conditions, Assumption (A1)(g) is usually satisfied. There also exist simple methods to determine the mixing rates for various classes of random processes, for example, Gaussian, Markov, autoregressive moving average, ARCH, and GARCH. Hence the assumption of a known mixing rate is reasonable and has been adopted in many studies, for example, Györfi, Härdle, Sarda, and Vieu (1989), Irle (1997), Meir (2000), Modha and Masry (1998), Roussas (1988), and Yu (1993). Auestad and Tjøstheim (1990) provided excellent discussions on the role of mixing for model identification in nonlinear time series analysis. But since the restriction of Assumption (A1)(c) as discussed before, ARCH or GARCH type processes may not satisfy all assumptions here. Finally conditions (A2)(d) and (e) are adopted from conditions of Lemma 3.2 of Yoshihara (1976), which are required to get the asymptotic equivalence of the nondegenerate U-statistic and its projection under the β-mixing process. They are technical assumptions and are quite standard.

LEMMA 3.1 (Franke and Mwita, 2003). *Suppose conditions (A1)(e)–(g) and (A2)(c) hold. The bandwidth sequence is such that* $a = \mathcal{O}(1)$ *and* $\tilde{S}_T = Ta^p(s_T \log T)^{-1} \to \infty$ *for some* $s_T \to \infty$. *Let* $S_T = a^2 + \tilde{S}_T^{-1/2}$. *Then for the nonparametric kernel estimator of the conditional quantile of* $\hat{Q}_\theta(x_t)$, *equation (9), we have*

$$\sup_{\|x\|\in G}\left|\hat{Q}_\theta(x) - Q_\theta(x)\right| = \mathcal{O}(S_T) + \mathcal{O}\left(\frac{1}{Ta^p}\right) \quad a.s. \tag{13}$$

LEMMA 3.2 (Li, 1999). *Let* $L_t = (\varepsilon_t, z_t)^T$ *be a stochastic process that satisfies conditions (A1)(a)–(d).* $\varepsilon_t \in R$, $z_t \in R^m$, *and* $K(\cdot)$ *be the kernel function with* h *being the smoothing parameter that satisfies conditions (A2)(a) and (b). Define*

$$\sigma_\varepsilon^2(z) = \mathrm{E}[\varepsilon_t^2 \mid z_t = z] \quad and \tag{14}$$

$$J_T \equiv \frac{1}{T(T-1)h^m}\sum_{t=1}^{T}\sum_{s\neq t}^{T} K_{ts}\varepsilon_t\varepsilon_s. \tag{15}$$

Then

$$Th^{m/2}J_T \;\to\; N(0,\sigma_0^2) \quad in\ distribution, \tag{16}$$

where $\sigma_0^2 = 2\mathrm{E}\left\{\sigma_\varepsilon^4(z_t)f_z(z_t)\right\}\left\{\int K^2(u)du\right\}$ *and* $f_z(\cdot)$ *is the marginal density function of* z_t.

We consider testing for local departures from the null that converge to the null at the rate $T^{-1/2}h^{-m/4}$. More precisely we consider the sequence of local alternatives

$$H_{1T} \;:\; F_{y|z}\{Q_\theta(x_t) + d_T l(z_t)|z_t\} = \theta, \tag{17}$$

where $d_T = T^{-1/2}h^{-m/4}$ and the function $l(\cdot)$ and its first-order derivatives are bounded.

THEOREM 3.1. *Assume the conditions (A1) and (A2). Then*

(i) Under the null hypothesis (3), $Th^{m/2}\hat{J}_T \xrightarrow{L} N(0,\sigma_0^2)$ *in distribution, where*

$$\sigma_0^2 = 2\mathrm{E}\left\{\sigma_\varepsilon^4(z_t)f_z(z_t)\right\}\left\{\int K^2(u)du\right\} \quad and$$

$$\sigma_\varepsilon^2(z_t) = \mathrm{E}(\varepsilon_t^2|z_t) = \theta(1-\theta).$$

(ii) Under the null hypothesis (3), $\hat{\sigma}_0^2 \equiv 2\theta^2(1-\theta)^2 1/(T(T-1)h^m)\sum_{s\neq t} K_{ts}^2$ *is a consistent estimator of* $\sigma_0^2 = 2\mathrm{E}\left\{\sigma_\varepsilon^4(z_t)f_z(z_t)\right\}\int K^2(u)du$. *Thus*

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

$$Th^{m/2}\hat{J}_T/\hat{\sigma}_0$$
$$= \sqrt{\frac{T}{T-1}}\frac{\sum_{t=1}^{T}\sum_{s\neq t}^{T} K_{ts}\left[1\left\{y_t \leqslant \hat{Q}_\theta(x_t)\right\}-\theta\right]\left[1\left\{y_s \leqslant \hat{Q}_\theta(x_s)\right\}-\theta\right]}{\sqrt{2}\theta(1-\theta)\sqrt{\sum_{s\neq t} K_{ts}^2}}.$$

(iii) Under the alternative hypothesis (4),

$$\hat{J}_T \to \mathrm{E}\{[F_{y|z}(Q_\theta(x_t)|z_t)-\theta]^2 f_z(z_t)\} > 0 \quad \textit{in probability.}$$

(iv) Under the local alternatives (A.2) in the Appendix, $Th^{m/2}\hat{J}_T \to N(\mu,\sigma_1^2)$ *in distribution, where*

$$\mu = \mathrm{E}\left[f_{y|z}^2\{Q_\theta(z_t)|z_t\}l^2(z_t)f_z(z_t)\right],$$
$$\sigma_1^2 = 2\mathrm{E}\left\{\sigma_v^4(z_t)f_z(z_t)\right\}\int K^2(u)du, \quad \textit{and}$$
$$\sigma_v^2(z_t) = \mathrm{E}(v_t^2|z_t) \quad \textit{with } v_t \equiv I\{y_t \leqslant Q_\theta(x_t)\} - F(Q_\theta(x_t)|z_t).$$

Theorem 3.1 generalizes the results of Zheng (1998) for independent data to the weakly dependent data case. A detailed proof of Theorem 3.1 is given in the Appendix. The main difficulty in deriving the asymptotic distribution of the statistic defined in equation (12) is that a nonparametric quantile estimator is included in the indicator function that is not differentiable with respect to the quantile argument and thus prevents taking a Taylor expansion around the true conditional quantile $Q_\theta(x_t)$. To circumvent the problem, Zheng (1998) made use of the work of Sherman (1994) on uniform convergence of U-statistics indexed by parameters. In this paper, we bound the test statistic by the statistics in which the nonparametric quantile estimator in the indicator function is replaced with sums of the true conditional quantile and upper and lower bounds consistent with the uniform convergence rate of the nonparametric quantile estimator, $1(y_t \leqslant Q_\theta(x_t) - C_T)$ and $1(y_t \leqslant Q_\theta(x_t) + C_T)$.

An important further step is to show that the differences of the ideal test statistic J_T given in equation (8) and the statistics having the indicator functions obtained from the first step stated previously are asymptotically negligible. We may directly show that the second moments of the differences are asymptotically negligible by using the result of Yoshihara (1976) on the bound of moments of U-statistics for absolutely regular processes. However, it is tedious to get bounds on the second moments with dependent data. In the proof we use instead the fact that the differences are second-order degenerate U-statistics. Thus by using the result on the asymptotic normal distribution of the second-order degenerate U-statistic of Fan and Li (1999), we can derive the asymptotic variance that is based on the independent and identically distributed (i.i.d.) sequence having the same marginal distributions as weakly dependent variables in the test statistic. With this little

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

trick we only need to show that the asymptotic variance is $\mathcal{O}(1)$ in an i.i.d. situation. For details refer to the Appendix.

4. SIMULATION

We generate bivariate data $\{y_t, w_t\}_{t=1}^T$ according to the following model:

$$y_t = \frac{1}{2} y_{t-1} + c w_{t-1}^2 + \varepsilon_{1t},$$
$$w_t = 1 + \frac{1}{2} w_{t-1} + \varepsilon_{2t},$$

where ε_{1t} and ε_{2t} are independent standard normal random variables. Here $c = 0$ corresponds to the hypothetical model; that is, w_t does not cause y_t in the θ quantile with respect to $\{y_{t-1}, w_{t-1}\}$. All the coefficients are set such that the corresponding time series is stationary and β-mixing with corresponding densities bounded to satisfy the assumptions discussed before. We use different values of $c \in [0,1]$ to investigate the power of the test, such that the higher c is, the stronger the causality of w_t on y_t is. Without loss of generality, we choose $\theta = 0.1, 0.5, 0.9$ and $T = 500, 1{,}000, 5{,}000$ here with the bandwidth h and a as in (7) and (10) as for a typical Nadaraya–Watson type estimator. We consider the nominal 0.05 significance level and repeat the test 500 times to generate the power.

Table 1 displays the power performance of the test for different combinations of T, c, and θ. First, obviously the power is very sensitive to the choice of T; that is, the larger T is, for the same c and θ, the larger the power is. From a technical point of view, this makes sense, because the more data we have, the more evidence we can draw from to detect the "causality" effect. Our asymptotic result, Theorem 3.1, needs the plug-in estimation of the asymptotic covariance matrix that is used to normalize the test statistic. Note that such an estimator is model-dependent and under the alternative is consistent with a different value than the one under the null. As a result, the power deteriorates for small T. On the other hand, whether the causality effect exists or not is the nature of the series, which is independent of the sample size used in this technical test. Enhancing the power performance for small-sample data using the simulation-based method deserves further research. Second, as discussed before, the higher c is, the stronger the causality of w_t on y_t is, which is confirmed by the larger and larger power values. Third, for different quantiles θ, we find that the powers with respect to $\theta = 0.5$ are usually larger than the powers with respect to $\theta = 0.1$ and 0.9.

5. APPLICATION TO COMMODITY PRICES

In financial and commodity markets, it has been argued that the covariation of the tails may be different from that of the rest of the distribution. The gold market is one of the most important markets in the world, where trading takes place 24 hours a day around the globe and transactions involving billions of dollars of

TABLE 1. Power performance for different combinations of T, c, and θ

c	Power (θ 0.1)	c	Power (θ 0.5)	c	Power (θ 0.9)
			$T = 500$		
0.00	0.024	0.00	0.108	0.00	0.010
0.03	0.030	0.03	0.288	0.03	0.020
0.06	0.058	0.06	0.796	0.06	0.108
0.09	0.190	0.09	0.991	0.09	0.585
0.12	0.414	0.12	1.000	0.12	0.950
0.15	0.696	0.15	1.000	0.15	0.994
0.18	0.888	0.18	1.000	0.18	1.000
0.21	0.962	0.21	1.000	0.21	1.000
0.24	0.988	0.24	1.000	0.24	1.000
0.27	1.000	0.27	1.000	0.27	1.000
0.30	1.000	0.30	1.000	0.30	1.000
			$T = 1{,}000$		
0.00	0.014	0.00	0.130	0.00	0.018
0.01	0.022	0.01	0.144	0.01	0.024
0.02	0.038	0.02	0.296	0.02	0.024
0.03	0.026	0.03	0.564	0.03	0.040
0.04	0.060	0.04	0.788	0.04	0.108
0.05	0.110	0.05	0.946	0.05	0.284
0.06	0.196	0.06	0.990	0.06	0.506
0.07	0.356	0.07	1.000	0.07	0.838
0.08	0.530	0.08	1.000	0.08	0.950
0.09	0.676	0.09	1.000	0.09	0.994
0.10	0.816	0.10	1.000	0.10	0.996
0.11	0.906	0.11	1.000	0.11	1.000
0.12	0.958	0.12	1.000	0.12	1.000
0.13	0.972	0.13	1.000	0.13	1.000
0.14	0.994	0.14	1.000	0.14	1.000
0.15	0.998	0.15	1.000	0.15	1.000
0.16	1.000	0.16	1.000	0.16	1.000
			$T = 5{,}000$		
0.00	0.020	0.00	0.116	0.00	0.026
0.01	0.028	0.01	0.328	0.01	0.046
0.02	0.124	0.02	0.904	0.02	0.142
0.03	0.490	0.03	1.000	0.03	0.728
0.04	0.924	0.04	1.000	0.04	0.988
0.05	1.000	0.05	1.000	0.05	1.000
0.06	1.000	0.06	1.000	0.06	1.000
0.07	1.000	0.07	1.000	0.07	1.000
0.08	1.000	0.08	1.000	0.08	1.000
0.09	1.000	0.09	1.000	0.09	1.000
0.10	1.000	0.10	1.000	0.10	1.000

TABLE 2. Unit root tests

Variable	Test type	Time trend term	Test statistics	CR value 5%	Unit root	Unit root after differencing
LN Oil	DF	no	0.86955	−1.94160	yes	no
	ADF	no	0.72255	−1.94160	yes	no
	PP	no	0.73107	−1.94160	yes	no
	KPSS	no	2.16221	0.14600	yes	no
	DF	include	−0.81819	−2.86386	yes	no
	ADF	include	−1.03287	−2.86386	yes	no
	PP	include	−0.94355	−2.86386	yes	no
	KPSS	include	2.16221	0.14600	yes	no
GBP	DF	no	−0.12461	−1.94160	yes	no
	ADF	no	−0.16186	−1.94160	yes	no
	PP	no	−0.12506	−1.94160	yes	no
	KPSS	no	5.26720	0.14600	yes	no
	DF	include	−1.53295	−2.86386	yes	no
	ADF	include	−1.51000	−2.86386	yes	no
	PP	include	−1.53853	−2.86386	yes	no
	KPSS	include	5.26720	0.14600	yes	no
LN Gold	DF	no	0.45931	−1.94160	yes	no
	ADF	no	1.03139	−1.94160	yes	no
	PP	no	0.69975	−1.94160	yes	no
	KPSS	no	3.50910	0.14600	yes	no
	DF	include	−1.98422	−2.86386	yes	no
	ADF	include	−1.36627	−2.86386	yes	no
	PP	include	−1.66336	−2.86386	yes	no
	KPSS	include	3.50910	0.14600	yes	no

Note: "LN Oil", "GBP", and "LN Gold" refer to the logarithmic Brent crude oil price, USD/GBP exchange rate, and logarithmic NYMEX spot gold price, respectively. The "Test types" DF, ADF, PP, and KPSS refer to unit root tests of, respectively, Dickey–Fuller (Fuller, 1976), augmented Dickey–Fuller (Fuller, 1976), Phillips–Perron (Phillips & Perron, 1988), and (Kwaitkowski et al., 1992).

gold are carried out each day. Understanding the mechanism of gold price changes is important for many outstanding issues in international economics and finance. Market participants are increasingly concerned with their exposure to large gold price fluctuations with special interest in which factors drive the changes. In this section, we apply the quantile causality test to investigate relations between the Brent crude oil, USD/GBP exchange rate and NYMEX spot gold prices (in USD per barrel and per ounce, respectively). The data, as seen in Figure 1, obtained from Datastream, are daily observations from 20 February 1997 to 17 July 2009 ($T = 3{,}237$). We use the USD/GBP instead of USD/EUR because the euro was only introduced as a new currency from 1 January 1999. As indicated by Table 2, we assume differenced logarithmic data are stationary and β-mixing with corresponding densities bounded. Because a long memory effect is not expected, we choose $p = q = 1$ and $m = 2$.

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

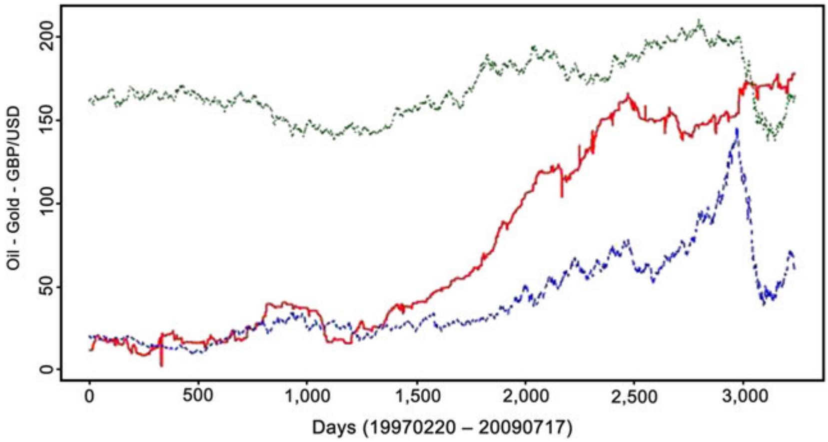

FIGURE 1. Plot of the gold prices, oil price, and exchange rate time series from 20 February 1997 to 17 July 2009.

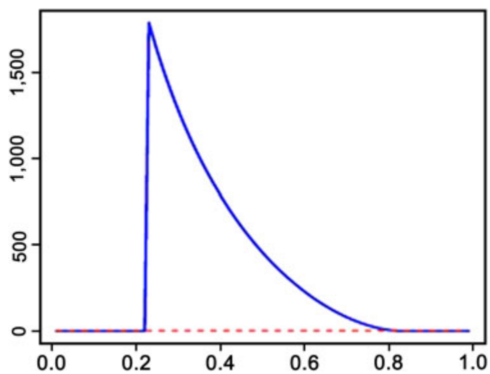

FIGURE 2. Test statistics with respect to different quantiles for the oil-gold prices causality test.

Figures 2 and 3 present results of testing whether oil prices Granger cause gold prices and whether the USD/GBP exchange rate Granger causes gold prices at the various quantiles, respectively, where logarithmic returns instead of the raw observations are used. The solid line and dotted line represent the standardized test statistics with respect to different quantiles (x-axis) and the critical value 1.96, respectively. In Figures 2 and 3, because the test statistic exceeds the critical value when $0.22 \le \theta \le 0.80$, we conclude that the oil price and exchange rate changes do not cause the gold price change in $\theta < 0.22$ or $\theta > 0.80$, whereas it is a prima facie cause in the $0.22 \le \theta \le 0.80$ quantile, respectively. For example, the oil price and USD/GBP exchange rate increases suggest that investors are wary of the U.S. dollar's weakness and inflation. Because gold is typically bought as an

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

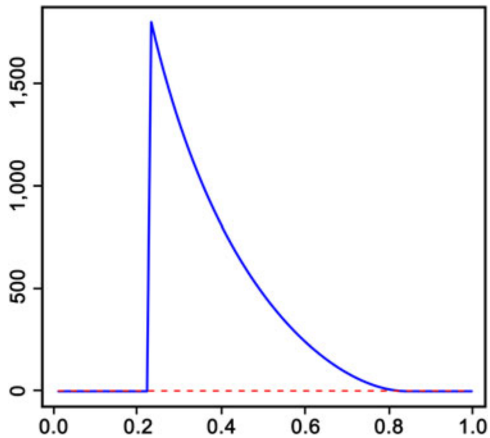

FIGURE 3. Test statistics with respect to different quantiles for the exchange rate-gold prices causality test.

alternative to the dollar among safe-haven assets, investors seeking safety from inflation risk and currency devaluation will cause the gold price to rise. However, the extreme low and high changes of the gold market may be caused by speculation. This is consistent with most of the empirical findings in the literature that the codependency may be stronger in the center than in the tails. By combining results from Figures 2 and 3, we find that the oil price and exchange rate changes have a significant predictive power for nonextreme gold price changes, which is, however, not significant for extreme changes. This finding could help to make it possible to use the gold price and GBP to hedge oil price changes in a more precise way with more careful investigation of their relations, which deserves further research.

6. CONCLUSION

By extending the Zheng (1998) idea to dependent data, we propose a consistent test for Granger causality in conditional quantile. The appealing feature of our proposed test is that it can investigate Granger causality in various conditional quantiles. The benefit of this is illustrated in the commodity market application where the causal relationships among the oil price, USD/GBP exchange rate, and gold price were shown to be different between a tail area and in the center of the distribution. We also illustrate that oil price and USD/GBP changes have significant predictive power on nonextreme gold price changes.

The test can be extended in a number of ways to test conditional quantile restrictions with dependent data: First, it can be extended to test functional misspecification, or the insignificance of a subset of regressors in quantile regression function, and second, it can also be used to test some semiparametric versus nonparametric models in quantile regression models.

REFERENCES

Auestad, B. & D. Tjøstheim (1990) Identification of nonlinear time series: First order characterisation and order determination. *Biometrica* 77, 669–687.

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

Bollerslev, T. (2001) Financial econometrics: Past developments and future challenges. *Journal of Econometrics* 100, 41–51.

Buchinsky, M. (1995) Quantile regression, Box-Cox transformation model, and the U.S. wage structure. 1963–1987. *Journal of Econometrics* 65, 65–154.

Campbell, J. & J. Cochrane (1999) By force of habit: A consumption-based explanation of aggregate stock market behaviour. *Journal of Political Economy* 107, 205–251.

Day, J.C. & E.C. Newburger (2002) The big payoff: Educational attainment and synthetic estimates of work-life earnings. Special studies. Current population reports. Statistical report p23–210, U.S. Department of Commerce, U.S. Census Bureau.

Diebolt, J. & D. Guégan (1993) Tail behavior of the stationary density of general nonlinear autoregressive processes of order 1. *Journal of Applied Probability* 30, 315–329.

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

Fan, Y. & Q. Li (1996) Consistent model specification tests: Omitted variables, parametric and semiparametric functional forms. *Econometrica* 64, 865–890.

Fan, Y. & Q. Li (1999) Central limit theorem for degenerate U-statistics of absolutely regular processes with applications to model specification tests. *Journal of Nonparametric Statistics* 10, 245–271.

Franke, J. & P. Mwita (2003) Nonparametric Estimates for Conditional Quantiles of Time Series. Wirtschaftsmathematik 87, University of Kaiserslautern.

Fuller, W. (1976) *Introduction to Statistical Time Series*. Wiley.

Granger, C. (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–438.

Granger, C. (1988) Some recent developments in a concept of causality. *Journal of Econometrics* 39, 199–211.

Györfi, L., W. Härdle, P. Sarda, & P. Vieu (1989) *Nonparametric Curve Estimation from Time Series*. Springer-Verlag.

Hall, P. (1984) Central limit theorem for integrated square error of multivariate nonparametric density estimators. *Journal of Multivariate Analysis* 14, 1–16.

Härdle, W., Y. Ritov, & S. Song (2009) Bootstrap Partial Linear Quantile Regression and Confidence Bands. SFB649 Discussion paper 2010-002, Humboldt Universität zu Berlin.

Härdle, W. & T. Stoker (1989) Investigating smooth multiple regression by the method of average derivatives. *Journal of the American Statistical Association* 84, 986–995.

Hong, Y., Y. Liu, & S. Wang (2009) Granger causality in risk and detection of extreme risk spillover between financial markets. *Journal of Econometrics* 150, 271–287.

Hsiao, C. & Q. Li (2001) A consistent test for conditional heteroskedasticity in time-series regression models. *Econometric Theory* 17, 188–221.

Irle, A. (1997) On the consistency in nonparametric estimation under mixing assumptions. *Multivariate Analysis* 60, 123–147.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, & Y. Shin (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54, 159–178.

Lee, T. & W. Yang (2007) Money-Income Granger-Causality in Quantiles. Manuscript, University of California, Riverside.

Li, Q. (1999) Consistent model specification tests for time series econometric models. *Journal of Econometrics* 92, 101–147.

Li, Q. & S. Wang (1998) A simple consistent bootstrap test for a parametric regression functional form. *Journal of Econometrics* 87, 145–165.

Masry, E. & D. Tjøstheim (1995) Nonparametric estimation and identification of nonlinear ARCH time series: Strong convergence and asymptotic normality. *Econometric Theory* 11, 258–289.
Masry, E. & D. Tjøstheim (1997) Additive nonlinear ARX time series and projection estimates. *Econometric Theory* 13, 214–252.
Meir, R. (2000) Nonparametric time series prediction through adaptive model selection. *Machine Learning* 39, 5–34.
Modha, D. & E. Masry (1998) Memory-universal prediction of stationary random processes. *IEEE Transactions on Information Theory* 44, 117–133.
Phillips, P.C.B. & P. Perron (1988) Testing for a unit root in time series regression. *Biometricka* 75, 335–346.
Powell, J., J. Stock, & T. Stoker (1989) Semiparametric estimation of index coefficients. *Econometrica* 57, 1403–1430.
Robinson, P.M. (1988) Root-n-consistent semiparametric regression. *Econometrica* 56, 931–954.
Roussas, G.G. (1988) Nonparametric estimation in mixing sequences of random variables. *Journal of Statistical Planning and Inference* 18, 135–149.
Sherman, R. (1994) Maximal inequalities for degenerate U-processes with applications to optimization estimators. *Annals of Statistics* 22, 439–459.
Yoshihara, K. (1976) Limiting behavior of u-statistics for stationary absolutely regular processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 35, 237–252.
Yu, B. (1993) Density estimation in the l1 norm for dependent data with applications. *Annals of Statistics* 21, 711–735.
Zheng, J. (1998) A consistent nonparametric test of parametric regression models under conditional quantile restrictions. *Econometric Theory* 14, 123–138.

APPENDIX

Proof of Theorem 3.1(i). In the proof, we use several approximations to $\hat{J}_T$. We define them now and recall a few already defined statistics for convenience of reference.

$$\hat{J}_T \equiv \frac{1}{T(T-1)h^m} \sum_{t=1}^{T} \sum_{s \neq t}^{T} K_{ts} \hat{\varepsilon}_t \hat{\varepsilon}_s, \tag{A.1}$$

$$J_T \equiv \frac{1}{T(T-1)h^m} \sum_{t=1}^{T} \sum_{s \neq t}^{T} K_{ts} \varepsilon_t \varepsilon_s, \tag{A.2}$$

$$J_{TU} \equiv \frac{1}{T(T-1)h^m} \sum_{t=1}^{T} \sum_{s \neq t}^{T} K_{ts} \varepsilon_{tU} \varepsilon_{sU}, \tag{A.3}$$

$$J_{TL} \equiv \frac{1}{T(T-1)h^m} \sum_{t=1}^{T} \sum_{s \neq t}^{T} K_{ts} \varepsilon_{tL} \varepsilon_{sL}, \tag{A.4}$$

where
$$\begin{aligned}
\hat{\varepsilon}_t &= I\left\{y_t \leqslant \hat{Q}_\theta(x_t)\right\} - \theta, \\
\varepsilon_t &= I\{y_t \leqslant Q_\theta(x_t)\} - \theta, \\
\varepsilon_{tU} &= I\{y_t + C_T \leqslant Q_\theta(x_t)\} - \theta, \\
\varepsilon_{tL} &= I\{y_t - C_T \leqslant Q_\theta(x_t)\} - \theta,
\end{aligned}$$

and C_T is an upper bound consistent with the uniform convergence rate of the nonparametric estimator of conditional quantile given in equation (13). The proof of Theorem 3.1(i) consists of three steps.

Step 1. Asymptotic normality.

$$Th^{m/2}J_T \to N(0,\sigma_0^2), \tag{A.5}$$

where $\sigma_0^2 = 2\mathrm{E}\left\{\theta^2(1-\theta)^2 f(z_t)\right\}\left\{\int K^2(u)du\right\}$ under the null.

Step 2. Conditional asymptotic equivalence. Suppose that both $Th^{m/2}(J_T - J_{TU}) = \mathcal{O}_p(1)$ and $Th^{m/2}(J_T - J_{TL}) = \mathcal{O}_p(1)$.

$$\text{Then} \quad Th^{m/2}(\hat{J}_T - J_T) = \mathcal{O}_p(1). \tag{A.6}$$

Step 3. Asymptotic equivalence.

$$Th^{m/2}(J_T - J_{TU}) = \mathcal{O}_p(1) \quad \text{and} \quad Th^{m/2}(J_T - J_{TL}) = \mathcal{O}_p(1). \tag{A.7}$$

The combination of steps 1–3 yields Theorem 3.1(i).

Proof of Step 1. Because J_T is a degenerate U-statistic of order 2, the result follows from Lemma 3.2. ■

Proof of Step 2. The proof of step 2 is motivated by the technique of Härdle and Stoker (1989) that was used in treating trimming an indicator function asymptotically. Suppose that the following two statements hold:

$$Th^{m/2}(J_T - J_{TU}) = \mathcal{O}_p(1) \quad \text{and} \tag{A.8}$$

$$Th^{m/2}(J_T - J_{TL}) = \mathcal{O}_p(1). \tag{A.9}$$

Use C_T to denote an upper bound consistent with the uniform convergence rate of the nonparametric estimator of conditional quantile given in equation (13). Suppose that

$$\sup|\hat{Q}_\theta(x) - Q_\theta(x)| \leqslant C_T. \tag{A.10}$$

If inequality (A.10) holds, then the following statements also hold:

$$\{Q_\theta(x) > y_t + C_T\} \subset \{\hat{Q}_\theta(x) > y_t\} \subset \{Q_\theta(x) > y_t - C_T\}, \tag{A.11}$$

$$1(Q_\theta(x) > y_t + C_T) \leqslant 1(\hat{Q}_\theta(x) > y_t) \leqslant 1(Q_\theta(x) > y_t - C_T), \tag{A.12}$$

$$J_{TU} \leqslant \hat{J}_T \leqslant J_{TL}, \tag{A.13}$$

$$|Th^{m/2}(J_T - \hat{J}_T)| \leqslant \max\{|Th^{m/2}(J_T - J_{TU})|, |Th^{m/2}(J_T - J_{TL})|\}. \tag{A.14}$$

Using (A.10) and (A.14), we have the following inequality:

$$\begin{aligned} &\mathrm{P}\left\{|Th^{m/2}(J_T - \hat{J}_T)| > \delta |\sup\left|\hat{Q}_\theta(x) - Q_\theta(x)\right| \leqslant C_T\right\} \\ &\leqslant \mathrm{P}\left\{\max\{|Th^{m/2}(J_T - J_{TU})|, |Th^{m/2}(J_T - J_{TL})|\} > \delta \middle| \sup|\hat{Q}_\theta(x) - Q_\theta(x)| \leqslant C_T\right\} \\ &\quad \text{for all } \delta > 0. \end{aligned} \tag{A.15}$$

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

Invoking Lemma 3.1 and condition (A2)(c), we have

$$\mathrm{P}\left\{\sup|\hat{Q}_\theta(x)-Q_\theta(x)|\leqslant C_T\right\} \to 1 \quad \text{as } T \to \infty. \tag{A.16}$$

By (A.8) and (A.9), as $T \to \infty$, we have

$$\mathrm{P}\left\{\max\{|Th^{m/2}(J_T-J_{TU})|,|Th^{m/2}(J_T-J_{TL})|\}>\delta\right\}\to 0 \quad \text{for all } \delta>0. \tag{A.17}$$

Therefore, as $T\to\infty$,

the right-hand side of the inequality (A.15) $\times$ $\mathrm{P}\left\{\sup|\hat{Q}_\theta(x)-Q_\theta(x)|\leqslant C_T\right\}\to 0$;

the left-hand side of the inequality (A.15) $\times$ $\mathrm{P}\left\{\sup|\hat{Q}_\theta(x)-Q_\theta(x)|\leqslant C_T\right\}$

$$=\mathrm{P}\left\{|Th^{m/2}(J_T-\hat{J}_T)|>\delta\right\}\to 0.$$

In summary, we have that if both $Th^{m/2}(J_T-J_{TU})=\mathcal{O}_p(1)$ and $Th^{m/2}(J_T-J_{TL})=\mathcal{O}_p(1)$, then $Th^{m/2}(\hat{J}_T-J_T)=\mathcal{O}_p(1)$. ■

Proof of Step 3. In the remaining proof, we focus on showing that $Th^{m/2}(J_T-J_{TU})=\mathcal{O}_p(1)$, with the proof of $Th^{m/2}(J_T-J_{TL})=\mathcal{O}_p(1)$ being treated similarly. The proof of step 3 is close in line with the proof in Zheng (1998). Denote

$$H_T(s,t,g)\equiv K_{ts}\{1(y_t\leqslant g(x_t))-\theta\}\{1(y_s\leqslant g(x_s))-\theta\} \quad \text{and} \tag{A.18}$$

$$J[g]\equiv\frac{1}{T(T-1)h^m}\sum_{t=1}^{T}\sum_{s\neq t}^{T}H_T(s,t,g). \tag{A.19}$$

Then we have $J_T\equiv J[Q_\theta]$ and $J_{TU}\equiv J[Q_\theta-C_T]$. We decompose $H_T(s,t,g)$ into three parts:

$$\begin{aligned}H_T(s,t,g)&=K_{ts}\{1(y_t\leqslant g(x_t))-F(g(x_t)|z_t)\}\{1(y_s\leqslant g(x_s))-F(g(x_s)|z_s)\}\\&\quad+2\times K_{ts}\{1(y_t\leqslant g(x_t))-F(g(x_t)|z_t)\}\{F(g(x_s)|z_s)-\theta\}\\&\quad+K_{ts}\{F(g(x_t)|z_t)-\theta\}\{F(g(x_s)|z_s)-\theta\}\\&=H_{1T}(s,t,g)+2H_{2T}(s,t,g)+H_{3T}(s,t,g).\end{aligned} \tag{A.20}$$

Then let $J_j[g]=1/(T(T-1)h^m)\sum_{t=1}^{T}\sum_{s\neq t}^{T}H_{jT}(s,t,g)$, $i=1,2,3$. We will treat $J_j[Q_\theta]-J_j[Q_\theta-C_T]$ for $j=1,2,3$ separately.

(1) $Th^{m/2}\left[J_1(Q_\theta)-J_1(Q_\theta-C_T)\right]=\mathcal{O}_p(1)$. By simple manipulation, we have

$$\begin{aligned}&J_1(Q_\theta)-J_1(Q_\theta-C_T)\\&\quad=\frac{1}{T(T-1)h^m}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\left[H_{1T}(s,t,Q_\theta)-H_{1T}(s,t,Q_\theta-C_T)\right]\\&\quad=\frac{1}{T(T-1)h^m}\sum_{t=1}^{T}\sum_{s\neq t}^{T}K_{ts}\Big\{[1(y_t\leqslant Q_\theta(x_t))-F(Q_\theta(x_t)|z_t)]\\&\qquad\qquad\times[1(y_s\leqslant Q_\theta(x_s))-F(Q_\theta(x_s)|z_s)]\end{aligned}$$

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

$$
\begin{aligned}
&-[1(y_t \leqslant (Q_\theta(x_t) - C_T)) - F((Q_\theta(x_t) - C_T)|z_t)] \\
&\times [1(y_s \leqslant (Q_\theta(x_s) - C_T)) - F((Q_\theta(x_s) - C_T)|z_s)] \Big\}.
\end{aligned} \tag{A.21}
$$

To avoid tedious work to get bounds on the second moment of $J_1(Q_\theta) - J_1(Q_\theta - C_T)$ with dependent data, we note that the right-hand side of (A.21) is a degenerate U-statistic of order 2. Thus we can apply Lemma 3.2 and have

$$
Th^{m/2}\left[J_1(Q_\theta) - J_1(Q_\theta - C_T)\right] \to N(0, \sigma_2^2) \quad \text{in distribution,} \tag{A.22}
$$

where the definition of the asymptotic variance σ_2^2 is based on the i.i.d. sequence having the same marginal distributions as weakly dependent variables in (A.21). That is,

$$
\sigma_2^2 = 2h^{-m}\tilde{\mathrm{E}}\left[H_{1T}(s,t,Q_\theta) - H_{1T}(s,t,Q_\theta - C_T)\right]^2,
$$

where the notation $\tilde{\mathrm{E}}$ is an expectation evaluated at an i.i.d. sequence having the same marginal distribution as the mixing sequences in (A.21) (Fan and Li, 1999, p. 248). Now, to show that $Th^{m/2}\left[J_1(Q_\theta) - J_1(Q_\theta - C_T)\right] = \mathcal{O}_p(1)$, we only need to show that the asymptotic variance $\sigma_2^2(z)$ is $\mathcal{O}(1)$ with i.i.d. data. Use Λ_T to denote an upper bound consistent with the integral over K_{ts} being of the order $\mathcal{O}(h^m)$. We have

$$
\begin{aligned}
&\tilde{\mathrm{E}}\left[H_{1T}(s,t,Q_\theta) - H_{1T}(s,t,Q_\theta - C_T)\right]^2 \\
&\quad \leqslant \Lambda_T \tilde{\mathrm{E}}\{[1_t(Q_\theta) - F_t(Q_\theta)][1_s(Q_\theta) - F_s(Q_\theta)] \\
&\qquad\qquad - [1_t(Q_\theta - C_T) - F_t(Q_\theta - C_T)][1_s(Q_\theta - C_T) - F_s(Q_\theta - C_T)]\}^2 \\
&\quad \leqslant \Lambda_T \tilde{\mathrm{E}}\{F_t(Q_\theta)[1 - F_t(Q_\theta)]F_s(Q_\theta)[1 - F_s(Q_\theta)]\} \\
&\qquad + \tilde{\mathrm{E}}\{F_t(Q_\theta - C_T)[1 - F_t(Q_\theta - C_T)]F_s(Q_\theta - C_T)[1 - F_s(Q_\theta - C_T)]\} \\
&\qquad - 2\mathrm{E}\{[F_t(\min(Q_\theta, Q_\theta - C_T)) - F_t(Q_\theta)F_t(Q_\theta - C_T)] \\
&\qquad\qquad \times [F_s(\min(Q_\theta, Q_\theta - C_T)) - F_s(Q_\theta)F_s(Q_\theta - C_T)]\} \\
&\quad = \Lambda_T \tilde{\mathrm{E}}\{[F_t(Q_\theta) - F_t(Q_\theta)F_t(Q_\theta)][F_s(Q_\theta) - F_s(Q_\theta)F_s(Q_\theta)]\} \\
&\qquad - \Lambda_T \tilde{\mathrm{E}}\{[F_t(\min(Q_\theta, Q_\theta - C_T)) - F_t(Q_\theta)F_t(Q_\theta - C_T)] \\
&\qquad\qquad \times [F_s(\min(Q_\theta, Q_\theta - C_T)) - F_s(Q_\theta)F_s(Q_\theta - C_T)]\} \\
&\qquad + \Lambda_T \tilde{\mathrm{E}}\{[F_t(Q_\theta - C_T) - F_t(Q_\theta - C_T)F_t(Q_\theta - C_T)] \\
&\qquad\qquad \times [F_s(Q_\theta - C_T) - F_s(Q_\theta - C_T)F_s(Q_\theta - C_T)]\} \\
&\qquad - \Lambda_T \tilde{\mathrm{E}}\{[F_t(\min(Q_\theta, Q_\theta - C_T)) - F_t(Q_\theta)F_t(Q_\theta - C_T)] \\
&\qquad\qquad \times [F_s(\min(Q_\theta, Q_\theta - C_T)) - F_s(Q_\theta)F_s(Q_\theta - C_T)]\} \\
&\quad \leqslant \Lambda_T C_T.
\end{aligned} \tag{A.23}
$$

Thus we have that $\sigma_2^2 = \mathcal{O}(C_T) = \mathcal{O}(1)$, and so

$$
Th^{m/2}\left[J_1(Q_\theta) - J_1(Q_\theta - C_T)\right] = \mathcal{O}_p(1). \tag{A.24}
$$

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

(2) $Th^{m/2}\left[J_2(Q_\theta)-J_2(Q_\theta-C_T)\right]=\mathcal{O}_p(1)$. Note that $H_{2T}(s,t,Q_\theta)=0$ because of $F_{y|z}(Q_\theta(x_s)|z_s)-\theta=0$. Then we have

$$\begin{aligned}J_2(Q_\theta)-J_2(Q_\theta-C_T)&=-J_2(Q_\theta-C_T)\\&=-\frac{1}{T(T-1)}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\frac{1}{h^m}K\left(\frac{z_t-z_s}{h}\right)\\&\quad\times\{1(y_t\leqslant Q_\theta(x_t)-C_T)-F_{y|z}(Q_\theta(x_t)-C_T|z_t)\}\\&\quad\times\{F_{y|z}(Q_\theta(x_s)-C_T|z_s)-\theta\}.\end{aligned}\tag{A.25}$$

By taking a Taylor expansion of $F_{y|z}(Q_\theta(x_s)-C_T|z_s)$ around $Q_\theta(x_s)$, it equals

$$\begin{aligned}&-\frac{1}{T(T-1)}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\frac{1}{h^m}K\left(\frac{z_t-z_s}{h}\right)\\&\times\{1(y_t\leqslant Q_\theta(x_t)-C_T)-F_{y|z}(Q_\theta(x_t)-C_T|z_t)\}\\&\times(-C_T)f_{y|z}(\bar{Q}_\theta(x_s)|z_s),\end{aligned}\tag{A.26}$$

where $\bar{Q}_\theta$ is between Q_θ and $Q_\theta-C_T$. Thus we have

$$\begin{aligned}&(J_2(Q_\theta)-J_2(Q_\theta-C_T))^2\\&\leqslant\left[\frac{1}{T(T-1)}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\frac{1}{h^m}K\left(\frac{z_t-z_s}{h}\right)\right.\\&\qquad\left.\times\left\{1(y_t\leqslant Q_\theta(x_t)-C_T)-F_{y|z}(Q_\theta(x_t)-C_T|z_t)\right\}\right]^2\Lambda^2C_T^2\\&=\Lambda^2C_T^2\left[\frac{1}{T}\sum_{t=1}^{T}\left\{1(y_t\leqslant Q_\theta(x_t)-C_T)-F_{y|z}(Q_\theta(x_t)-C_T)\right\}\hat{f}_z(z_t)\right]^2\\&\equiv\Lambda^2C_T^2\left\{\frac{1}{T}\sum_{t=1}^{T}u_t\hat{f}_z(z_t)\right\}^2\\&=\Lambda^2C_T^2T^{-2}\sum_{t=1}^{T}u_t^2\hat{f}_z^2(z_t)+\Lambda^2C_T^2T^{-2}\sum_{t=1}^{T}\sum_{s\neq t}^{T}u_tu_s\hat{f}_z(z_t)\hat{f}_z(z_s)\\&\equiv J_{21}+J_{22},\end{aligned}\tag{A.27}$$

where the inequality holds because of Assumption (A.1)(e).

$$\begin{aligned}\mathrm{E}|J_{21}|&=\Lambda^2C_T^2T^{-1}\left[T^{-1}\sum_{t=1}^{T}\mathrm{E}\left\{u_t^2\hat{f}_z^2(z_t)\right\}\right]\\&=O\left(C_T^2T^{-2}h^{-m}\right),\end{aligned}\tag{A.28}$$

where the second equality is derived by using Lemma C.3(iii) of Li (1999).

$$\begin{aligned} J_{22} = \Lambda^2 C_T^2 \Bigg[& T^{-2} \sum_{t=1}^{T} \sum_{s \neq t}^{T} u_t u_s f_z(z_t) f_z(z_s) \\ & + 2T^{-2} \sum_{t=1}^{T} \sum_{s \neq t}^{T} u_t u_s f_z(z_t) \left\{ \hat{f}_z(z_s) - f_z(z_s) \right\} \\ & + T^{-2} \sum_{t=1}^{T} \sum_{s \neq t}^{T} u_t u_s \left\{ \hat{f}_z(z_t) - f_z(z_t) \right\} \left\{ \hat{f}_z(z_s) - f_z(z_s) \right\} \Bigg] \\ \equiv \Lambda C_T^2 & (J_{221} + J_{222} + J_{223}). \end{aligned} \quad \textbf{(A.29)}$$

Following the line of the proof of Lemma A.2(i) of Li (1999) we have that

$$J_{221} = \mathcal{O}_p\left(T^{-2}\right), \qquad J_{222} = \mathcal{O}_p\left(T^{-1}\right), \quad \text{and} \quad J_{223} = \mathcal{O}_p\left(T^{-1}\right); \quad \text{thus}$$
$$J_{22} = \mathcal{O}_p\left(C_T^2 T^{-1}\right). \quad \textbf{(A.30)}$$

Thus, combining (A.28) and (A.30), we have

$$\begin{aligned} Th^{m/2}\left[J_2(Q_\theta) - J_2(Q_\theta - C_T)\right] &= \mathcal{O}_p(C_T) + \mathcal{O}_p\left(C_T T^{1/2} h^{m/2}\right) \\ &= \mathcal{O}_p(1). \end{aligned} \quad \textbf{(A.31)}$$

(3) $Th^{m/2}\left[J_3(Q_\theta) - J_3(Q_\theta - C_T)\right] = \mathcal{O}_p(1)$. Noting that $H_{3T}(s,t,Q_\theta) = 0$ because of $F(Q_\theta(x_j)|z_j) - \theta = 0$ for $j = t, s$, we have

$$\begin{aligned} & J_3(Q_\theta) - J_3(Q_\theta - C_T) \\ & \quad = -\frac{1}{T(T-1)} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \frac{1}{h^m} K\left(\frac{z_t - z_s}{h}\right) \\ & \qquad \times \{F(Q_\theta(x_t) - C_T|z_t) - \theta\}\{F(Q_\theta(x_s) - C_T|z_s) - \theta\} \\ & \quad = \frac{1}{T(T-1)} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \frac{1}{h^m} K\left(\frac{z_t - z_s}{h}\right) C_T^2 f_{y|z}(\bar{Q}_\theta(x_t)|z_t) f_{y|z}(\bar{Q}_\theta(x_s)|z_s) \\ & \quad = C_T^2 \frac{1}{T} \sum_{t=1}^{T} f_{y|z}(\bar{Q}_\theta(x_t)|z_t) f_{y|z}(\bar{Q}_\theta(x_s)|z_s) \hat{f}_z(z_t). \end{aligned} \quad \textbf{(A.32)}$$

Thus, we have

$$\begin{aligned} & \mathrm{E}\left|J_3(Q_\theta) - J_3(Q_\theta - C_T)\right| \\ & \quad \leqslant \Lambda C_T^2 \frac{1}{T} \sum_{t=1}^{T} \mathrm{E}\left|\hat{f}_z(z_t)\right| \\ & \quad \leqslant \Lambda C_T^2 \frac{1}{T} \sum_{t=1}^{T} \mathrm{E}\left|f_z(z_t)\right| + \Lambda C_T^2 \frac{1}{T} \sum_{t=1}^{T} \mathrm{E}\left|\hat{f}_z(z_t) - f_z(z_t)\right| \\ & \quad = \mathcal{O}\left(C_T^2\right). \end{aligned} \quad \textbf{(A.33)}$$

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

Finally, we have

$$Th^{m/2}\left[J_3(Q_\theta)-J_3(Q_\theta-C_T)\right]=\mathcal{O}_p\left(Th^{m/2}C_T^2\right)$$
$$=\mathcal{O}_p(1). \tag{A.34}$$

By combining (A.24), (A.31), and (A.34), we have the result of step 3. ■

Proof of Theorem 3.1(ii). Because

$$\sigma_0^2=2\theta^2(1-\theta)^2\mathrm{E}\{f_z(z_t)\}\int K^2(u)du \quad \text{and}$$
$$\hat{\sigma}_0^2\equiv 2\theta^2(1-\theta)^2\frac{1}{T(T-1)h^m}\sum_{s\neq t}K_{ts}^2,$$

it is enough to show that

$$\sigma_T^2\equiv\frac{1}{T(T-1)h^m}\sum_{s\neq t}K_{ts}^2$$
$$=\mathrm{E}\{f_z(z_t)\}\int K^2(u)du+\mathcal{O}_p(1). \tag{A.35}$$

Note that σ_T^2 is a nondegenerate U-statistic of order 2 with kernel

$$H_T(z_t,z_s)=\frac{1}{h^m}K^2\left(\frac{z_t-z_s}{h}\right). \tag{A.36}$$

Because Assumptions (A2)(d) and (e) satisfy the conditions of Lemma 3.2 of Yoshihara (1976) on the asymptotic equivalence of the U-statistic and its projection under β-mixing, we have for $\gamma=2(\delta-\delta')/\delta'(2+\delta)>0$

$$\sigma_T^2\equiv\frac{1}{T(T-1)}\sum_{s\neq t}H_T(z_t,z_s)$$
$$=\int\int H_T(z_1,z_2)dF_z(z_1)dF_z(z_2)$$
$$+2T^{-1}\sum_{t=1}^{T}\left[\int H_T(z_t,z_2)dF_z(z_2)-\int\int H_T(z_1,z_2)dF_z(z_1)dF_z(z_2)\right]$$
$$+\mathcal{O}_p(T^{-1-\gamma})$$
$$=\int\int H_T(z_1,z_2)dF_z(z_1)dF_z(z_2)+\mathcal{O}_p(1)$$
$$=\int\int\frac{1}{h^m}K^2\left(\frac{z_1-z_2}{h}\right)dF_z(z_1)dF_z(z_2)+\mathcal{O}_p(1)$$
$$=\int K^2(u)du\int f_z^2(z)dz+\mathcal{O}_p(1). \tag{A.37}$$

The result of Theorem 3.1(ii) follows from (A.37). ■

Proof of Theorem 3.1(iii). The proof of Theorem 3.1(iii) consists of two steps.

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

Step 1. Show that $\hat{J}_T = J_T + \mathcal{O}_p(1)$ under the alternative hypothesis (4).

Step 2. Show that $J_T = J + \mathcal{O}_p(1)$ under the alternative hypothesis (4), where $J = \mathrm{E}\{[F_{y|z}(Q_\theta(x_t)|z_t) - \theta]^2 f_z(z_t)\}$. The combination of steps 1 and 2 yields Theorem 3.1(iii).

Proof of Step 1. We note that the results of step 2 and $Th^{m/2}\left[J_1(Q_\theta) - J_1(Q_\theta - C_T)\right] = \mathcal{O}_p(1)$ of step 3 in the proof of Theorem 3.1(i) still hold under the alternative hypothesis (4). Thus we focus on showing that $J_2(Q_\theta) - J_2(Q_\theta - C_T) = \mathcal{O}_p(1)$ and $J_3(Q_\theta) - J_3(Q_\theta - C_T) = \mathcal{O}_p(1)$.

We begin with showing that $J_2(Q_\theta) - J_2(Q_\theta - C_T) = \mathcal{O}_p(1)$. By the same procedures as in (A.27), we can show that $J_2(Q_\theta - C_T) = \mathcal{O}_p(T^{-1}h^{-m/2})$. Thus it remains to show that $J_2(Q_\theta) = \mathcal{O}_p(1)$. By taking a Taylor expansion of $F_{y|z}(Q_\theta(x_s)|z_s)$ around $Q_\theta(x_s)$, we have

$$\begin{aligned} J_2(Q_\theta) &= -\frac{1}{T(T-1)} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \frac{1}{h^m} K\left(\frac{z_t - z_s}{h}\right) \\ &\quad \times \{1(y_t \leqslant Q_\theta(x_t)) - F_{y|z}(Q_\theta(x_t)|z_t)\} \times f_{y|z}(\bar{Q}_\theta(x_s)|z_s) \\ &= \frac{1}{T} \sum_{t=1}^{T} \{1(y_t \leqslant Q_\theta(x_t)) - F_{y|z}(Q_\theta(x_t))\} f_{y|z}(\bar{Q}_\theta(x_s)|z_s) \hat{f}_z(z_t) \\ &\equiv \frac{1}{T} \sum_{t=1}^{T} u_t f_{y|z}(\bar{Q}_\theta(x_s)|z_s) \hat{f}_z(z_t). \end{aligned} \tag{A.38}$$

By similar arguments as in (A.26) and (A.31), we have

$$J_2(Q_\theta) = \mathcal{O}\left(T^{-1}h^{-m}\right). \tag{A.39}$$

Next, we show that $Th^{m/2}\left[J_3(Q_\theta) - J_3(Q_\theta - C_T)\right] = \mathcal{O}_p(1)$ under the alternative hypothesis (4). Because $F(Q_\theta(x_j)|z_j) - \theta \neq 0$ for $j = t, s$ under the alternative hypothesis, we have

$$\begin{aligned} &J_3(Q_\theta) - J_3(Q_\theta - C_T) \\ &\quad = \frac{1}{T(T-1)} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \frac{1}{h^m} K\left(\frac{z_t - z_s}{h}\right) \{F(Q_\theta(x_t)|z_t) - \theta\}\{F(Q_\theta(x_s)|z_s) - \theta\} \\ &\qquad - \frac{1}{T(T-1)} \sum_{t=1}^{T} \sum_{s \neq t}^{T} \frac{1}{h^m} K\left(\frac{z_t - z_s}{h}\right) \\ &\qquad \times \{F(Q_\theta(x_t) - C_T|z_t) - \theta\}\{F(Q_\theta(x_s) - C_T|z_s) - \theta\} \\ &\quad = \frac{1}{T} \sum_{t=1}^{T} \{F(Q_\theta(x_t)|z_t) - \theta\}\{F(Q_\theta(x_s)|z_s) - \theta\} \hat{f}_z(z_t) \\ &\qquad - \frac{1}{T} \sum_{t=1}^{T} \{F(Q_\theta(x_t) - C_T|z_t) - \theta\}\{F(Q_\theta(x_s) - C_T|z_s) - \theta\} \hat{f}_z(z_t). \end{aligned} \tag{A.40}$$

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

By taking a Taylor expansion of $F_{y|z}(Q_\theta(x_j) - C_T|z_j)$ around $Q_\theta(z_j)$ for $j = t, s$, we have

$$J_3(Q_\theta) - J_3(Q_\theta - C_T) = \frac{1}{T}\sum_{t=1}^{T}\{F(Q_\theta(x_t)|z_t) - \theta\}C_T f_{y|z}(\bar{Q}_\theta(x_t)|z_t)\hat{f}_z(z_t)$$
$$+\frac{1}{T}\sum_{t=1}^{T} C_T f_{y|z}(\bar{Q}_\theta(x_t)|z_t)\{F(Q_\theta(x_s)|z_s) - \theta\}\hat{f}_z(z_t)$$
$$-\frac{1}{T}\sum_{t=1}^{T} C_T^2 f_{y|z}(\bar{Q}_\theta(x_t)|z_t) f_{y|z}(\bar{Q}_\theta(x_s)|z_s)\hat{f}_z(z_t). \qquad \textbf{(A.41)}$$

We further take a Taylor expansion of $F_{y|z}(Q_\theta(x_j)|z_j)$ around $Q_\theta(z_j)$ for $j = t, s$ and have

$$J_3(Q_\theta) - J_3(Q_\theta - C_T) = \frac{1}{T}\sum_{t=1}^{T} f_{y|z}(\bar{Q}_\theta(x_t, z_t)|z_t)C_T f_{y|z}(\bar{Q}_\theta(x_s)|z_s)\hat{f}_z(z_t)$$
$$+\frac{1}{T}\sum_{t=1}^{T} C_T f_{y|z}(\bar{Q}_\theta(x_t)|z_t) f_{y|z}(\bar{Q}_\theta(x_s, z_s)|z_s)\hat{f}_z(z_t)$$
$$-\frac{1}{T}\sum_{t=1}^{T} C_T^2 f_{y|z}(\bar{Q}_\theta(x_t)|z_t) f_{y|z}(\bar{Q}_\theta(x_s)|z_s)\hat{f}_z(z_t), \qquad \textbf{(A.42)}$$

where $\bar{Q}_\theta(x_s, z_s)$ is between $Q_\theta(x_s)$ and $Q_\theta(z_s)$. Then by using the same procedures as in (A.30), we have

$$J_3(Q_\theta) - J_3(Q_\theta - C_T) = \mathcal{O}(C_T). \qquad \textbf{(A.43)}$$

Now we have the result of step 1 for the proof of Theorem 3.1(iii). ■

Proof of Step 2. Using (7) and the uniform convergence rate of the kernel regression estimator under a β-mixing process, we have

$$J_T = \frac{1}{T(T-1)h^m}\sum_{t=1}^{T}\sum_{s\neq t}^{T} K_{ts}\varepsilon_t\varepsilon_s$$
$$= \frac{1}{T}\sum_{t=1} \hat{\mathrm{E}}(\varepsilon_t|z_t)\hat{f}_z(z_t)\varepsilon_t$$
$$= \frac{1}{T}\sum_{t=1} \mathrm{E}(\varepsilon_t|z_t) f_z(z_t)\varepsilon_t + \frac{1}{T}\sum_{t=1}\left\{\hat{\mathrm{E}}(\varepsilon_t|z_t)\hat{f}_z(z_t) - \mathrm{E}(\varepsilon_t|z_t) f_z(z_t)\right\}\varepsilon_t$$
$$= \frac{1}{T}\sum_{t=1} \mathrm{E}(\varepsilon_t|z_t) f_z(z_t)\varepsilon_t + \mathcal{O}_p(1)$$
$$= \mathrm{E}\left[\mathrm{E}(\varepsilon_t|z_t) f_z(z_t)\varepsilon_t\right] + \mathcal{O}_p(1)$$
$$= J + \mathcal{O}_p(1). \qquad \textbf{(A.44)}$$

■

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

Proof of Theorem 3.1(iv). The proof of Theorem 3.1(iv) is close in line with the proof in Zheng (1998). The proof of Theorem 3.1(iv) consists of two steps.

Step 1. Show that $\hat{J}_T = J_T + \mathcal{O}_p(T^{-1}h^{-m/2})$ under the alternative hypothesis (A.2).

Step 2. Show that $Th^{m/2}J_T \to N(\mu, \sigma_1^2)$ under the alternative hypothesis (A.2), where $\mu = \mathrm{E}\left[f_{y|z}^2\{Q_\theta(z_t)|z_t\}l^2(z_t)f_z(z_t)\right]$, $\sigma_1^2 = 2\mathrm{E}\left\{\sigma_\upsilon^4(z_t)f_z(z_t)\right\}$ $\int K^2(u)du$, and $\sigma_\upsilon^2(z_t) = \mathrm{E}(\upsilon_t^2|z_t)$ with $\upsilon_t \equiv I\{y_t \leqslant Q_\theta(x_t)\} - F(Q_\theta(x_t)|z_t)$.

Proof of Step 1. The results of step 1 in the proof of Theorem 3.1(iii) show that, under the general alternative hypothesis (4), the elements consisting of $\hat{J}_T - J_T$ are all $\mathcal{O}_p(T^{-1}h^{-m/2})$ except for $J_2(Q_\theta(x))$, the order of which is $\mathcal{O}\left(T^{-1}h^{-m}\right)$ as in (A.39). Thus we need to show that $J_2(Q_\theta(x)) = \mathcal{O}_p(T^{-1}h^{-m/2})$ under the local alternative hypothesis (A.2). Taking a Taylor expansion of $F_{y|z}\{Q_\theta(z_t) + d_T l(z_t)|z_t\}$ around $d_T = 0$, we have

$$F_{y|z}\{Q_\theta(z_t) + d_T l(z_t)|z_t\} = \theta + d_T f_{y|z}\{Q_\theta(z_t)|z_t\}l(z_t) + \mathcal{O}_p(d_T^2). \tag{A.45}$$

By similar procedures as in (A.38) and (A.39), we have

$$\begin{aligned}
J_2(Q_\theta(x)) &= -\frac{1}{T(T-1)}\sum_{t=1}^{T}\sum_{s\neq t}^{T}\frac{1}{h^m}K\left(\frac{z_t - z_s}{h}\right)\{1(y_t \leqslant Q_\theta(x_t)) - F_{y|z}(Q_\theta(x_t)|z_t)\} \\
&\quad \times d_T f_{y|z}\{Q_\theta(z_t)|z_t\}l(z_t) + \mathcal{O}_p\left(d_T^2\right) \\
&= -d_T\frac{1}{T}\sum_{t=1}^{T}\{1(y_t \leqslant Q_\theta(x_t)) - F_{y|z}(Q_\theta(x_t)|z_t)\} \\
&\quad \times f_{y|z}\{Q_\theta(z_t)|z_t\}l(z_t)\hat{f}_z(z_t) + \mathcal{O}_p\left(d_T^2\right) \\
&\equiv -d_T\frac{1}{T}\sum_{t=1}^{T}u_t f_{y|z}\{Q_\theta(z_t)|z_t\}l(z_t)\hat{f}_z(z_t) + \mathcal{O}_p\left(d_T^2\right) \\
&= \mathcal{O}_p\left(d_T^2\right).
\end{aligned} \tag{A.46}$$

■

Proof of Step 2. Taking a Taylor expansion of $F_{y|z}\{Q_\theta(z_t) + d_T l(z_t)|z_t\}$ around $d_T = 0$, we have

$$\begin{aligned}
J_T(Q_\theta(x)) &= \frac{1}{T(T-1)h^m}\sum_{t=1}^{T}\sum_{s\neq t}^{T}K_{ts}\{1(y_t \leqslant Q_\theta(x_t))F(Q_\theta(x_t)|z_t)\} \\
&\quad \times \{1(y_s \leqslant Q_\theta(x_s)) - F(Q_\theta(x_s)|z_s)\} \\
&\quad - \frac{2d_T}{T(T-1)h^m}\sum_{t=1}^{T}\sum_{s\neq t}^{T}K_{ts}\{1(y_t \leqslant Q_\theta(x_t)) - F(Q_\theta(x_t)|z_t)\} \\
&\quad \times f_{y|z}\{Q_\theta(z_s)|z_s\}l(z_s)
\end{aligned}$$

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press

$$+\frac{d_T^2}{T(T-1)h^m}\sum_{t=1}^{T}\sum_{s\neq t}^{T} K_{ts} f_{y|z}\{Q_\theta(z_t)|z_t\} l(z_t) f_{y|z}\{Q_\theta(z_s)|z_s\} l(z_s)$$

$$+\mathcal{O}_p\left(d_T^2\right)$$

$$= T_{1T} - 2d_T T_{2T} + d_T^2 T_{3T} + \mathcal{O}_p\left(d_T^2\right). \quad \textbf{(A.47)}$$

Noting that T_{1T} is a degenerate U-statistic of order 2, by Lemma 3.2, we have

$$Th^{m/2} T_{1T} \to N\left(0, \sigma_1^2\right) \quad \text{in distribution,} \quad \textbf{(A.48)}$$

Similarly to the proof for (A.31), we can show that $T_{2T} = \mathcal{O}\left\{(Th^m)^{-1}\right\}$, and so $d_T T_{2T} = \mathcal{O}\left\{(Th^{m/2})^{-1}\right\}$. And by the same procedures as in (A.44), we have

$$T_{3T} \to \mathrm{E}\left[f_{y|z}^2\{Q_\theta(z_t)|z_t\} l^2(z_t) f_z(z_t)\right] \quad \text{in probability.} \quad \textbf{(A.49)}$$

Thus,

$$Th^{m/2} J_T \to N\left(\mu, \sigma_1^2\right), \quad \textbf{(A.50)}$$

where $\mu = \mathrm{E}\left[f_{y|z}^2\{Q_\theta(z_t)|z_t\} l^2(z_t) f_z(z_t)\right]$. ■

https://doi.org/10.1017/S0266466611000685 Published online by Cambridge University Press