SPECIFICATION TESTS OF PARAMETRIC DYNAMIC CONDITIONAL QUANTILES*

J. Carlos Escanciano† — *Indiana University*

Carlos Velasco — *Universidad Carlos III de Madrid*

August 27, 2006

Abstract

This article proposes omnibus consistent goodness-of-fit tests of a parametric dynamic quantile regression model. Contrary to the existing procedures we allow for the simultaneous specification of an infinite number of quantiles under fairly weak conditions on the underlying data generating process. We study the asymptotic distribution of the test statistics under the null and under fixed and local alternatives. It turns out that the asymptotic null distribution depends on the data generating process and the hypothesized model. We propose and justify theoretically a subsampling procedure for approximating the asymptotic critical values of tests. The article also considers asymptotically distribution-free tests for the classical *location-scale* family based on certain weighted standardized residuals processes. An appealing property of all tests proposed in the article is that they do not require estimation of the non-parametric (conditional) sparsity function. A Monte Carlo study compares the proposed tests and shows that the asymptotic results provide good approximations for small sample sizes. Finally, an application of our methodology to the Sharpe Style Analysis of the Magellan Fund and a reanalysis of the Pennsylvania Reemployment Bonus Experiments provides evidence that the linear quantile model is a good specification for the first and a misspecified model for the second.

Keywords and Phrases: Omnibus tests; Conditional quantiles; Nonlinear time series; Empirical processes; Quantile processes; Subsampling; Distribution-free tests; Location-scale models.

---

*We would like to thank Professor Roger Koenker for helpful comments. Research funded by the Spanish Ministerio de Educación y Ciencia, reference number SEJ2004-04583/ECON.

†Parts of this paper were written while J. Carlos Escanciano was visiting the Cowles Foundation at Yale University and the Economics Department at Cornell University, whose hospitality is gratefully acknowledged. This author thanks Professor Peter C.B. Phillips and Professor Yongmiao Hong the possibility of these visits. Research funded by the Spanish Ministerio de Educación y Ciencia, reference number SEJ2005-07657/ECON.

1. INTRODUCTION

Quantile regression is a powerful alternative to least squares regression in a wide range of econometric applications that vary from labor economics or demand analysis to finance, see the special issue of Empirical Economics (2001, vol.26) and the references therein. The conditional quantile has the advantage over its natural competitor, the conditional mean, of being more robust to outliers and imposing less restrictions on the data generating process (DGP). Rather than relying on a single measure of conditional location, the quantile regression approach allows the researcher to explore a range of conditional quantile functions, thereby providing a more complete analysis of the conditional dependence structure of the variables under consideration. Since the seminal work by Koenker and Basset (1978) there has been a large body of research devoted to regression quantiles, resulting in a well-developed theory of asymptotic inference for many important aspects of quantile regression. Most of the extant literature has been devoted to the estimation of quantile parameters and the associated so-called quantile processes, see, e.g., Koenker and Xiao (2002). It is well-known that such inference procedures depend crucially on the validity of the specified parametric functional forms for the range of quantiles under consideration (cf. Kim and White, 2002). The main purpose of this article is to develop omnibus diagnostic tests for the correct specification of the functional form of a family of parametric conditional quantiles over a range of quantiles of interest and under fairly general conditions on the underlying DGP.

More precisely, let us consider the real-valued dependent variable Y_t, and the explanatory vector $I_{t-1} \in \mathbb{R}^d$, $d \in \mathbb{N}$, say. To be more concrete, let $Z_t \in \mathbb{R}^m$, $m \in \mathbb{N}$, be an m-dimensional observable random variable (r.v) and $W_{t-1} = (Y_{t-1}, ..., Y_{t-s})' \in \mathbb{R}^s$, where A' denotes the matrix transpose of A. The conditioning variable we consider is $I_{t-1} = (W'_{t-1}, Z'_t)'$, so $d = s+m$. We assume throughout the article that the time series process $\{(Y_t, Z'_t)' : t = 0, \pm 1, \pm 2, ...\}$, defined on the probability space $(\Omega, \mathcal{A}, P)$, is strictly stationary and ergodic. Assuming that the conditional distribution of Y_t given I_{t-1} is continuous, we define the α-th conditional quantile of Y_t given $I_{t-1} = x$ as the measurable function $q_\alpha(x)$ satisfying the equation

$$P(Y_t \leq q_\alpha(I_{t-1}) \mid I_{t-1}) = \alpha, \text{ almost surely (a.s.).} \quad (1)$$

In *parametric* quantile regression modeling one assumes the existence of a parametric family of functions $\mathcal{M} = \{m(\cdot, \theta(\alpha)) : \theta(\cdot) : \mathcal{T} \to \Theta \subset \mathbb{R}^p\}$, where $\mathcal{T} = [\epsilon, 1-\epsilon]$ is the range of quantiles of interest, with $\epsilon \in (0, 1/2]$, and one proceeds to make inference on $\theta(\cdot)$ or to test if $q_\cdot \in \mathcal{M}$, i.e., if there exists some $\theta_0(\cdot) : \mathcal{T} \to \Theta \subset \mathbb{R}^p$ such that $m(I_{t-1}, \theta_0(\alpha)) = q_\alpha(I_{t-1})$ a.s. $\forall \alpha \in \mathcal{T}$. We remark that our theory is also valid for a general compact set $\mathcal{T}$ of $(0, 1)$, but in accordance with the quantile regression literature we present our theory with $\mathcal{T} = [\epsilon, 1-\epsilon]$, $\epsilon \in (0, 1/2]$.

Leading examples of specifications $\mathcal{M}$ are the Linear Quantile Regression (LQR) model

$$m(I_{t-1}, \theta_0(\alpha)) \equiv m(Z_t, \theta_0(\alpha)) = Z_t' \theta_0(\alpha), \qquad \alpha \in \mathcal{T},$$

with the *location-scale* shift model as the prominent example in which $\theta_0(\alpha) = (\beta_0, \gamma_0 F_0^{-1}(\alpha)) \in \Theta \subset \mathbb{R}^p$, and where $F_0^{-1}(\alpha)$ denotes a univariate quantile function, see, e.g., Koenker and Xiao (2002), or the Linear Quantile Autoregression model of order s (LQAR(s)),

$$m(I_{t-1}, \theta_0(\alpha)) \equiv m(W_{t-1}, \theta_0(\alpha)) = \theta_{01}(\alpha) + W_{t-1}' \theta_{02}(\alpha), \qquad \theta_0(\alpha) = (\theta_{01}(\alpha), \theta_{02}'(\alpha))',$$

which arises, for instance, from the random coefficient model

$$Y_t = \theta_{01}(U_t) + W_{t-1}' \theta_{02}(U_t), \tag{2}$$

where $\theta_{01}(\cdot)$ and $\theta_{02}(\cdot)$ are such that the right hand side of (2) is monotone increasing in U_t, and $\{U_t\}$ are independent and identically distributed (*iid*) standard uniform random variables, see Koenker and Xiao (2004) for inferences on the LQAR(s) model.

Although much effort has been devoted to inferences on $\theta_0(\alpha)$ based on the associated quantile processes, i.e., $Q_n(\alpha) := \sqrt{n}\,(\theta_n(\alpha) - \theta_0(\alpha))$, for $\theta_n(\alpha)$ a $\sqrt{n}$-consistent estimator of $\theta_0(\alpha)$, and inferences based on $Q_n(\alpha)$ usually depend on the correct specification of the parametric regression quantile model, no consistent test for $q_\cdot \in \mathcal{M}$ has been proposed. In the present article we propose omnibus consistent tests for $q_\cdot \in \mathcal{M}$ valid for general linear and nonlinear quantile models under time series sequences.

The condition $q_\cdot \in \mathcal{M}$ can be equivalently expressed as an *infinite* number of conditional moment restrictions (CMR)

$$E[1(Y_t \leq m(I_{t-1}, \theta_0(\alpha))) - \alpha \mid I_{t-1}] = 0 \text{ a.s. for some } \theta_0(\cdot) : \mathcal{T} \to \Theta \subset \mathbb{R}^p, \forall \alpha \in \mathcal{T}. \tag{3}$$

Therefore, all of the many procedures available in the literature for testing a CMR can be applied for testing the correct specification of the parametric dynamic quantiles, with the proviso that an infinite number of CMR have to be tested. The vast amount of literature on testing CMR can be divided into two approaches. The first approach is called the "local approach", because it is based on nonparametric estimators of the conditional moment. Using this idea Zheng (1998) has proposed a quantile regression specification test based on kernel smoothing estimators of the conditional moment $E[1(Y_t \leq m(I_{t-1}, \theta_0(\alpha))) - \alpha \mid I_{t-1}]$ under *iid* observations for a fixed $\alpha \in (0, 1)$. Horowitz and Spokoiny (2002) have developed a specification test for LQR for the median function (i.e., $\alpha = 0.5$) which is uniformly consistent against smooth alternatives whose distance from the linear model converges to zero at the fastest possible rate, but the rate is slower than the parametric rate. Recently, Whang (2005) using ideas from the empirical likelihood literature has proposed a specification test for quantile regression and censored quantile regression for *iid* data. Local-based tests usually have known asymptotic null distributions after an appropriate choice of the bandwidth sequence, but they are not consistent against Pitman's local alternatives.

The second methodology in the CMR literature is called the "integrated approach", see Bierens (1982) and Stute (1997). Using this methodology, Bierens and Ginther (2001) proposed a test for (3) for a specific quantile, i.e., for a particular $\alpha \in (0,1)$. Their test is consistent against $n^{-1/2}$ local alternatives, with n the sample size, but it relies on an upper bound on the asymptotic critical value, which might be too conservative. Bierens and Ginther (2001) considered *iid* observations and do not take into account the uncertainty due to parameter estimation, see also Inoue (1999) for a related approach. Koul and Stute (1999) considered asymptotic pivotal tests for parametric conditional quantiles of first-order autoregressive processes. To obtain the pivotal property of the test they use a martingale transform (cf. Khmaladze, 1981). Alternatively, Whang (2004) has considered a subsampling approach to approximate the asymptotical critical values for multivariate LQR. Also recently, He and Zhu (2003) use empirical process theory to develop a bootstrap-based test for linear and nonlinear quantile regressions in an *iid* framework.

An important limitation for our purposes of all the aforementioned proposals is that they do not consider the problem (3), but the less restrictive problem of testing for $q_{\alpha_0} \in \mathcal{M}_{\alpha_0}$ for a fixed $\alpha_0 \in (0,1)$ and a parametric family $\mathcal{M}_{\alpha_0} = \{m(\cdot, \theta(\alpha_0)) : \theta(\alpha_0) \in \Theta \subset \mathbb{R}^p\}$. Unlike these procedures, our new tests consider the problem (3) for the whole set of quantiles of interest $\mathcal{T}$. The proposed tests are based on functionals of a quantile-marked empirical process. The asymptotic theory for the test statistics is derived using new weak convergence results for empirical processes under martingale conditions, which are of independent interest. It turns out that the asymptotic null distributions of test statistics depend on the specification under the null and the DGP. We propose to implement the test with the assistance of the subsampling. Another important contribution of the paper is the development of asymptotically distribution-free (ADF) tests based on a weighted standardized residual empirical process for testing the adequacy of the quantile regression model imposed by the classical *location-scale* model.

The rest of the article is organized as follows. In Section 2 we introduce the quantile-marked empirical process, which is the basis upon which the new test statistics for testing (3) are developed. We study the asymptotic distribution of the proposed tests under the null and under fixed and local alternatives. In Section 3 a subsampling procedure for approximating the asymptotic null distribution of the proposed omnibus tests is considered and theoretically justified. Section 4 is devoted to obtain ADF test statistics for the *location-scale* model. In Section 5 we make a simulation exercise comparing the subsampling and ADF tests under the null and under some alternatives. Finally, an application of our methodology to the Sharpe Style Analysis of the Magellan Fund and a reanalysis of the Pennsylvania Reemployment Bonus Experiments highlights the merits of our approach. Proofs are deferred to an appendix. Throughout the article A^c and $|A|$ denote the complex conjugate and Euclidean norm of A, respectively. In the sequel C is a generic constant that may change from one expression to another. The symbol $O_P(1)$ denotes boundedness in probability and $o_P(1)$ convergence to zero in probability. All limits are taken as the sample size $n \to \infty$.

2. TEST STATISTICS AND ASYMPTOTIC THEORY

The main goal of this article is to test the null hypothesis

$$H_0 : E[\Psi_\alpha(Y_t - m(I_{t-1}, \theta_0)) \mid I_{t-1}] = 0 \text{ a.s. for some } \theta_0 \in \mathcal{B} \text{ and for all } \alpha \in \mathcal{T},$$

against the nonparametric alternatives

$$H_A : P(E[\Psi_\alpha(Y_t - m(I_{t-1}, \theta(\alpha))) \mid I_{t-1}] \neq 0) > 0, \text{ for some } \alpha \in \mathcal{T} \text{ and for all } \theta(\alpha) \in \Theta \subset \mathbb{R}^p,$$

where $\Psi_\alpha(\varepsilon) = 1(\varepsilon \leq 0) - \alpha$, and $\mathcal{B}$ is a family of uniformly bounded functions from $\mathcal{T}$ to $\Theta \subset \mathbb{R}^p$. Note that under H_0 (and a mild continuity condition), $m(x, \theta_0(\alpha))$ is identified as the α-th quantile of the conditional distribution of Y_t given $I_{t-1} = x$, for all $\alpha \in \mathcal{T}$. Testing for H_0 is a challenging testing problem since it involves an infinite number of non-smooth CMR parametrized by $\alpha \in \mathcal{T}$. We address these technical difficulties by means of new weak convergence theorems for empirical process under martingale conditions, see the Appendix.

Using the results in Bierens (1982), our first aim is to characterize H_0 by the infinite number of unconditional moment restrictions

$$E[\Psi_\alpha(Y_1 - m(I_0, \theta_0)) \exp(ix' I_0)] = 0, \ \forall x \in \Upsilon \subset \mathbb{R}^d, \text{ for some } \theta_0 \in \mathcal{B} \text{ and for all } \alpha \in \mathcal{T}, \tag{4}$$

where Υ is a compact subset of $\mathbb{R}^d$ containing the origin, and $i = \sqrt{-1}$ is the imaginary unit. Instead of the exponential function we may also use any of the parametric families considered in Bierens and Ploberger (1997), see also Stinchcombe and White (1998).

Given a sample $\{(Y_t, I_{t-1}')' : 1 \leq t \leq n\}$ and a parameter value $\theta \in \mathcal{B}$, we consider the quantile-marked empirical process indexed by $x \in \Upsilon$, $\alpha \in \mathcal{T}$ and $\theta \in \mathcal{B}$,

$$S_n(x, \alpha, \theta) := n^{-1/2} \sum_{t=1}^{n} \Psi_\alpha(Y_t - m(I_{t-1}, \theta)) \exp(ix' I_{t-1}).$$

Associated to S_n are the quantile-marked *error* and *residual* processes, respectively, defined by $R_n(x, \alpha) \equiv S_n(x, \alpha, \theta_0)$ and $R_n^1(x, \alpha) \equiv S_n(x, \alpha, \theta_n)$, for a $\sqrt{n}-$consistent estimator $\theta_n(\alpha)$ of $\theta_0(\alpha)$, say. The null hypothesis is likely to hold when the process $R_n^1(x, \alpha)$ is close to zero for almost all $(x', \alpha)' \in \Pi := \Upsilon \times \mathcal{T}$.

The most popular estimator of θ_0 is the Quantile Regression Estimator (QRE), initially proposed Koenker and Basset (1978) for the linear model, and subsequently generalized to other frameworks by numerous authors, see references below. The QRE is defined as any solution $\theta_{KB,n}(\alpha)$ minimizing

$$\beta \longmapsto \sum_{t=1}^{n} \rho_\alpha(Y_t - m(I_{t-1}, \beta))$$

with respect to $\beta \in \Theta \subset \mathbb{R}^p$, where $\rho_\alpha(\varepsilon) = -\Psi_\alpha(\varepsilon)\,\varepsilon$. Koenker and Park (1996) discussed the existence of $\theta_{KB,n}(\alpha)$ and an interior point algorithm for its computation.

Basset and Koenker (1978) proved the consistency and asymptotic normality of $\theta_{KB,n}(\alpha)$ in the Linear Regression (LR) model, including the least absolute deviation estimator, see also Pollard (1991). The asymptotic theory for $\theta_{KB,n}$ based on the associated quantile process $Q_n(\cdot) = \sqrt{n}(\theta_{KB,n}(\cdot) - \theta_0(\cdot))$, as a process with parameter $\alpha \in \mathcal{T}$, have been considered, among others, in Gutenbrunner and Jurečkova (1992) and Gutenbrunner, Jurečkova, Koenker and Portnoy (1993) for LR models, in Koul and Saleh (1994) and Jurečkova and Hallin (1999) for linear autoregressions, and by Mukherjee (1999) for nonlinear autoregressions (NLAR). For early contributions see Portnoy (1984). In the present article we do not restrict ourselves to $\theta_{KB,n}$ and we consider any estimator θ_n satisfying some mild conditions, see A3 below.

The process R_n^1 is a mapping from $(\Omega, \mathcal{A}, P)$ with values in $\ell^\infty(\Pi)$, where $\ell^\infty(\Pi)$ is the space of all complex-valued functions that are uniformly bounded on Π. The space $\ell^\infty(\Pi)$ is furnished with the supremum metric, say d_∞, and let $\mathcal{B}_{d_\infty}$ be the corresponding Borel σ-algebra. Let $\Longrightarrow$ denote weak convergence on $(\ell^\infty(\Pi), \mathcal{B}_{d_\infty})$ in the sense of J. Hoffmann-Jørgensen, see, e.g., Dudley (1999, p. 94), or Definition 1.3.3 in van der Vaart and Wellner (1996).

Because of (4), test statistics are based on a distance from the standardized sample analogue of $E[\Psi_\alpha(Y_1 - m(I_0, \theta_0(\alpha))) \exp(ix'I_0)]$ to zero, i.e., on a norm of R_n^1, say $\Gamma(R_n^1)$. A popular norm is the Cramér-von Mises (CvM) functional

$$CvM_n := \int_{\Pi} \left| R_n^1(x, \alpha) \right|^2 d\Phi(x) dW(\alpha), \tag{5}$$

where Φ and W are some integrating measures on Υ and $\mathcal{T}$, respectively. Other continuous (with respect to d_∞) functionals Γ from $\ell^\infty(\Pi)$ to $\mathbb{R}$ are of course possible. Then, the omnibus tests we proposed in this article reject the null hypothesis H_0 for "large" values of $\Gamma(R_n^1)$. Practicalities about the test statistic CvM_n are discussed in Section 5.

2.1 Asymptotic null distribution.

In this subsection we establish the limit distribution of the quantile-marked empirical process R_n^1 under the null hypothesis H_0. The null limit distributions of the tests are the limit distributions of some continuous functionals of R_n^1. To derive asymptotic results we consider the following notation and assumptions. Throughout the paper the family $\mathcal{B}$, in which the parameter θ_0 takes values, is endowed with the sup norm, i.e., $\|\theta\|_{\mathcal{B}} = \sup_{\alpha \in \mathcal{T}} |\theta(\alpha)|$. Let for each $t \in \mathbb{Z}$, $\mathcal{F}_t = \sigma(I_t', I_{t-1}', ...)$, be the σ-field generated by the information set obtained up to time t. Let us define for each $t \in \mathbb{Z}$, the quantile innovation $\varepsilon_{t,\alpha} := Y_t - q_\alpha(I_{t-1})$ and the parametric quantile error $e_t(\theta(\alpha)) := Y_t - m(I_{t-1}, \theta(\alpha))$. Define also the family of conditional distributions

$$F_x(y) := P(Y_t \leq y \mid I_{t-1} = x), \qquad F_{x,\alpha}(y) := P(\varepsilon_{t,\alpha} \leq y \mid I_{t-1} = x). \tag{6}$$

Let $f_{I_0,\alpha}$ be the error density function of the cumulative distribution function (cdf) $F_{I_0,\alpha}$. Let $N_{[\cdot]}(\delta, \mathcal{H}, \|\cdot\|)$ be the δ-bracketing number of a class of functions $\mathcal{H}$ with respect to a norm $\|\cdot\|$, i.e., the smallest number r such that there exist $f_1, ..., f_r$ and $\Delta_1, ..., \Delta_r$ such that $\max_{1 \leq i \leq r} \|\Delta_i\| < \delta$ and for all $f \in \mathcal{H}$, there exists an $1 \leq i \leq r$ such that $\|f - f_i\| < \Delta_i$, see Definition 2.1.6 in van der Vaart and Wellner (1996).

Assumption A1:

A1(a): $\{(Y_t, Z_t')' : t = 0, \pm 1, \pm 2, ...\}$ is a strictly stationary and erdogic process and $(\Psi_\alpha(\varepsilon_{t,\alpha}), \mathcal{F}_{t-1})_{t \in \mathbb{Z}}$ is a martingale difference sequence for all $\alpha \in \mathcal{T}$.

A1(b): The parametric family $m(x, \theta_0(\alpha))$ is nondecreasing in α, $\forall x \in \mathbb{R}^d$.

A1(c): $E[|I_0|^2] < C$.

A1(d): The family of distributions functions $\{F_x, x \in \mathbb{R}^d\}$ has Lebesgue densities $\{f_x, x \in \mathbb{R}^d\}$ that are uniformly bounded

$$\sup_{x \in \mathbb{R}^d, y \in \mathbb{R}} |f_x(y)| \leq C$$

and equicontinuous: for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\sup_{x \in \mathbb{R}^d, |y-z| \leq \delta} |f_x(y) - f_x(z)| \leq \epsilon.$$

Assumption A2: For each $\theta_1 \in \mathcal{B}$,

A2(a): There exists a vector of functions $g : \mathbb{R}^p \times \Theta \to \mathbb{R}^q$ such that $g(I_{t-1}, \theta_1(\alpha))$ is $\mathcal{F}_{t-1}$-measurable for each $t \in \mathbb{Z}$, and satisfies, for all $k < \infty$,

$$\sup_{1 \leq t \leq n, \|\theta_1 - \theta_2\|_{\mathcal{B}} \leq k n^{-1/2}} n^{1/2} \|m(I_{t-1}, \theta_2) - m(I_{t-1}, \theta_1) - (\theta_2 - \theta_1)' g(I_{t-1}, \theta_1)\|_{\mathcal{B}} = o_P(1)$$

A2(b): For a sufficiently small $\delta > 0$,

$$E\left[\sup_{\|\theta_1 - \theta_2\|_{\mathcal{B}} \leq \delta} |1(Y_t \leq m(I_{t-1}, \theta_1(\alpha))) - 1(Y_t \leq m(I_{t-1}, \theta_2(\alpha)))|\right] \leq C\delta, \ \forall \alpha \in \mathcal{T} \text{ and}$$

$$E[\sup_{|\alpha_1 - \alpha_2| \leq \delta} |m(I_0, \theta_1(\alpha_1)) - m(I_0, \theta_1(\alpha_2))|] \leq C\delta.$$

A2(c): Uniformly in $\alpha \in \mathcal{T}$, $E\left|g\left(I_0, \theta_1(\alpha)\right)\right|^2 < \infty$, and uniformly in $(x', \alpha)' \in \Pi$,

$$\left|\frac{1}{n}\sum_{t=1}^{n} g(I_{t-1}, \theta_0(\alpha)) \exp(ix' I_{t-1}) f_{I_{t-1},\alpha}(0) - E\left[g(I_{t-1}, \theta_0(\alpha)) \exp(ix' I_{t-1}) f_{I_{t-1},\alpha}(0)\right]\right| = o_P(1).$$

Assumption A3:

A3(a): The parametric space Θ is compact in $\mathbb{R}^p$. The true parameter $\theta_0(\alpha)$ belongs to the interior of Θ for each $\alpha \in \mathcal{T}$, and $\theta_0 \in \mathcal{B}$. The class $\mathcal{B}$ satisfies

$$\int_0^\infty \left(\log(N_{[\cdot]}(\delta^2, \mathcal{B}, \|\cdot\|_{\mathcal{B}}))\right)^{1/2} d\delta < \infty.$$

A3(b): The estimator $\theta_n \in \mathcal{B}$, for all n sufficiently large, and satisfies the following asymptotic expansion under H_0 uniformly in $\alpha \in \mathcal{T}$,

$$Q_n(\alpha) = \sqrt{n}(\theta_n(\alpha) - \theta_0(\alpha)) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} l_\alpha(Y_t, I_{t-1}, \theta_0(\alpha)) + o_P(1),$$

where $l_\alpha(\cdot)$ is such that $E[l_\alpha(Y_1, I_0, \theta_0(\alpha))] = 0$, $L_\alpha(\theta_0(\alpha)) = E[l_\alpha(Y_1, I_0, \theta_0(\alpha)) l'_\alpha(Y_1, I_0, \theta_0(\alpha))]$ exists and is positive definite, and $E[l_\alpha(Y_t, I_{t-1}, \theta_0(\alpha)) \Psi_\alpha(Y_s - m(I_{s-1}, \theta_0(\alpha)))] = 0$ if $t \neq s$. Furthermore, as a process in $\ell^\infty(\mathcal{T})$, $Q_n(\alpha)$ converges weakly to a Gaussian process $Q(\cdot)$ with zero mean and covariance function

$$K_Q(\alpha_1, \alpha_2) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{s=1}^{n} E[l_{\alpha_1}(Y_t, I_{t-1}, \theta_0(\alpha_1)) l_{\alpha_2}(Y_s, I_{s-1}, \theta_0(\alpha_2))].$$

Assumption A1(a) is standard in the model checks literature under time series, see, e.g., Bierens and Ploberger (1997). A1(b) is natural in the present context. A1(c) is necessary for the equicontinuity of the limit process of R_n and can be avoided using $\exp(ix'\phi(I_{t-1}))$, with $\phi(\cdot)$ a one-to-one bounded mapping (see Bierens and Ginther, 2001), instead of $\exp(ix' I_{t-1})$. A1(d) is necessary for the tightness of the process R_n^1 and is required in Koul and Stute (1997). Assumptions A2(a)-A2(c) are classical in inference about nonlinear models, see Koul (2002) monograph. A2 is satisfied for all models considered in the literature under mild moment assumptions, e.g. LQR and LQAR models. Conditions for the satisfaction of A3(a) can be found in van der Vaart and Wellner (1996), see e.g. their Theorem 2.7.5 for monotone classes of functions which applies to LQAR models. The condition $\theta_n \in \mathcal{B}$, for all n sufficiently large, can be weakened to $P(\theta_n \in \mathcal{B}) \to 1$ as $n \to \infty$, at the cost of complicating the proofs, see Escanciano and Song (2006). A3(b) has been established in the literature under a variety of conditions and different models and DGP's, see, for instance, Theorem 1 in Gutenbrunner and Jurečkova (1992) or Theorem 3.2 in Mukherjee (1999). For NLAR models with *iid* innovations $(\varepsilon_t)_{t\in\mathbb{Z}}$ distributed as F_ε, Mukherjee (1999) proved A3 for $\theta_{KB,n}(\alpha)$. Then, under some mild additional assumptions, including that $\Sigma_{\theta_0(\alpha)} := E\left[g\left(I_1, \theta_0(\alpha)\right) g\left(I_1, \theta_0(\alpha)\right)'\right]$ exists and is positive definite, Mukherjee (1999) showed that A3(b) holds for the QRE under H_0 with

$$l_\alpha(Y_t, I_{t-1}, \theta_0(\alpha)) = -\frac{\Sigma^{-1}_{\theta_0(\alpha)} g(I_{t-1}, \theta_0(\alpha)) \Psi_\alpha(\varepsilon_t)}{q(\alpha)},$$

where $q(\alpha) = f_\varepsilon(F_\varepsilon^{-1}(\alpha))$ and f_ε is the density of F_ε. The quantile limit process $Q(\cdot)$ in that case is $\Sigma_{\theta_0(\cdot)}^{-1} W(\cdot)/q(\cdot)$, where $W(\cdot)$ denotes a vector of p independent Brownian bridges on $\mathcal{T}$.

Now, we establish the limit process of R_n. Under A1(a) and H_0, because $R_n(v)$ is a zero-mean square-integrable martingale for each $v = (x', \alpha)' \in \Pi$, using a suitable Central Limit Theorem (CLT) for stationary ergodic martingale difference sequences, cf. Billingsley (1961), we have that the finite-dimensional distributions of R_n converge to those of a multivariate normal distribution with a zero mean vector and variance-covariance matrix given by the covariance function

$$K_\infty(v_1, v_2) = (\alpha_1 \wedge \alpha_2 - \alpha_1\alpha_2) E[\exp(i(x_1 - x_2)' I_{t-1})], \tag{7}$$

where from now on $v_1 = (x_1', \alpha_1)'$ and $v_2 = (x_2', \alpha_2)'$ represent generic elements of Π, and $\wedge$ denotes the minimum, i.e., $a \wedge b = \min\{a, b\}$. The next result is an extension of the convergence of the finite-dimensional distributions of R_n to weak convergence in the space $\ell^\infty(\Pi)$.

THEOREM 1: *Under the null hypothesis H_0 and Assumptions A1(a-c)*

$$R_n \Longrightarrow R_\infty,$$

where R_∞ is a Gaussian process with zero mean and covariance function (7).

Theorem 1 generalizes Bierens and Ginther (2001) to a time series setup and more importantly, to the case in which all the quantiles in $\mathcal{T}$ are considered in the specification test. In other words, we consider the process R_n indexed by $x \in \Upsilon$ and $\alpha \in \mathcal{T}$, whereas their process is indexed only in $x \in \Upsilon$. Note that no mixing conditions are required in Theorem 1.

In practice, θ_0 is unknown and has to be estimated from a sample $\{(Y_t, I_{t-1}')' : 1 \leq t \leq n\}$ by an estimator θ_n. When we replace θ_0 in R_n by θ_n, resulting in R_n^1, we need to investigate how the estimation error will affect the asymptotic properties of R_n^1. The next result shows this effect on the asymptotic null distribution of R_n^1. Define the function

$$G(x, \theta_0(\alpha)) := E[g(I_0, \theta_0(\alpha)) f_{I_0,\alpha}(0) \exp(ix' I_0)], \qquad x \in \Upsilon,\ \alpha \in \mathcal{T}.$$

THEOREM 2: *Under the null hypothesis H_0 and Assumptions A1-A3*

$$\sup_{x \in \Upsilon, \alpha \in \mathcal{T}} \left| R_n^1(x, \alpha) - R_n(x, \alpha) + G'(x, \theta_0(\alpha)) n^{-1/2} \sum_{t=1}^{n} l_\alpha(Y_t, I_{t-1}, \theta_0(\alpha)) \right| = o_P(1).$$

As a consequence, we obtain the following corollary.

COROLLARY 1: *Under the assumptions of Theorem 2*

$$R_n^1 \Longrightarrow R_\infty^1,$$

where $R^1_\infty(\cdot) = R_\infty(\cdot) - G'(\cdot, \theta_0(\cdot))Q(\cdot)$ *(in distribution).*

Now, using the last corollary and the Continuous Mapping Theorem (CMT) we obtain the asymptotic null distribution of continuous functionals such as CvM_n.

Corollary 2: *Under the assumptions of Theorem 2, for any continuous functional* $\Gamma(\cdot)$ *from* $\ell^\infty(\Pi)$ *to* $\mathbb{R}$,

$$\Gamma(R^1_n) \xrightarrow{d} \Gamma(R^1_\infty).$$

2.2 Consistency and Pitman's local alternatives.

In this section we study the consistency properties of tests based on functionals $\Gamma(R^1_n)$. First, we show that these tests are consistent, that is, they are able to detect all alternatives in H_A. To that end, we need the following assumption.

Assumption A4: Under H_A there exists a $\theta_1 \in \mathcal{B}$ such that $\|\theta_n - \theta_1\|_{\mathcal{B}} = o_P(1)$.

See Kim and White (2003) for conditions on $\theta_{KB,n}$ to satisfy Assumption A4, see also Section 3 in Angrist, Chernozhukov and Fernández-Val (2006). Henceforth, almost sure convergence of nonmesurable maps is understood, as usual, as outer almost sure convergence, see van der Vaart and Wellner (1996) for definitions.

Theorem 3: *Under the alternative hypothesis* H_A *and Assumptions A1, A2, A3(a) and A4,*

$$n^{-1/2} R^1_n(\cdot) \xrightarrow{a.s} E[\Psi_\cdot(e_t(\theta_1(\cdot))) \exp(i \cdot I_{t-1})].$$

Furthermore, the function $E[\Psi_\cdot(e_t(\theta_1(\cdot))) \exp(i \cdot I_{t-1})]$ *is different from zero in a subset with positive Lebesgue measure on* Π.

A consequence of Theorem 3 and the CMT is that (under the assumptions of Theorem 3),

$$\int_\Pi \left| n^{-1/2} R^1_n(x, \alpha) \right|^2 d\Phi(x) dW(\alpha) \xrightarrow{P} \int_\Pi \left| E[\Psi_\alpha(e_t(\theta_1(\alpha))) \exp(ix' I_{t-1})] \right|^2 d\Phi(x) dW(\alpha) > 0,$$

provided that Φ and W are absolute continuous with respect to the Lebesgue measure. In such a situation, the test statistic CvM_n will diverge to $+\infty$ under any fixed alternative and the test will be consistent.

Now we analyse the asymptotic distribution of R^1_n under a sequence of local alternatives converging to null at a parametric rate $n^{-1/2}$. We consider the DGP generating the local alternatives

$$H_{A,n} : E[\Psi_\alpha(Y_t - m(I_{t-1}, \theta_0)) \mid I_{t-1}] = \frac{a_\alpha(I_{t-1})}{n^{1/2}} \text{ a.s. for some } \theta_0 \in \mathcal{B} \text{ and for all } \alpha \in \mathcal{T}, \quad (8)$$

where the function $a_\alpha(\cdot) : \mathbb{R}^d \longrightarrow \mathbb{R}$ satisfies the following assumption.

Assumption A5: $a_\alpha(\cdot)$ is such that $E \sup_{\alpha \in \mathcal{T}} |a_\alpha(I_{t-1})| < \infty$. There exists a $\mathcal{F}_{t-1}$-measurable r.v. C_{t-1} with $E[C_{t-1}^2] < \infty$, such that for all $t \in \mathbb{Z}$ and for all $\alpha_1, \alpha_2 \in \mathcal{T}$,

$$|a_{\alpha_1}(I_{t-1}) - a_{\alpha_2}(I_{t-1})| \leq C_{t-1} |\alpha_1 - \alpha_2|, \text{ a.s.}$$

To derive the next result we need the following assumption on the behaviour of the estimator under the local alternatives.

Assumption A3': The estimator $\theta_n(\alpha)$ satisfies the following asymptotic expansion under $H_{A,n}$, uniformly in α,

$$\sqrt{n}(\theta_n(\alpha) - \theta_0(\alpha)) = \xi_a(\alpha) + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} l_\alpha(Y_t, I_{t-1}, \theta_0(\alpha)) + o_P(1),$$

where the function $l_\alpha(\cdot)$ is as in A3(b) and $\xi_a(\alpha) \in \mathbb{R}^p$ for each $\alpha \in \mathcal{T}$.

Assumption A3' holds for most estimators considered in the literature. For instance, in the nonlinear time series context of Mukherjee (1999), the corresponding $\xi_a(\alpha)$ to $\theta_{KB,n}(\alpha)$ is

$$\xi_a(\alpha) = -q^{-1}(\alpha)\Sigma_{\theta_0(\alpha)}^{-1} E[f_{I_0,\alpha}(0) g(I_{t-1}, \theta_0) a_\alpha(I_{t-1})].$$

The shift in charge of local power against alternatives in $H_{A,n}$ is given by

$$D_a(x, \theta_0(\alpha), \alpha) := E[a_\alpha(I_0) \exp(ix' I_0)] - \xi_a'(\alpha) G(x, \theta_0(\alpha)).$$

THEOREM 4: *Under the local alternatives (8), Assumptions A1-A2, A3(a), A5 and A3'*

$$R_n^1 \Longrightarrow R_\infty^1 + D_a,$$

where R_∞^1 is the process defined in Theorem 2.

It is not difficult to show that

$$D_a \equiv 0 \text{ a.e.} \Longleftrightarrow a_\alpha(I_{t-1}) = \xi_a'(\alpha) g(I_{t-1}, \theta_0(\alpha)) \text{ for all } \alpha \in \mathcal{T} \text{ a.s.}$$

Therefore, for directions $a_\alpha(\cdot)$ not collinear to the score $g(\cdot, \theta_0(\alpha))$, the shift function D_a is non-trivial and test statistics based on $\Gamma(R_n^1)$ for a symmetric functional Γ are asymptotically strictly unbiased against the local alternatives (8). The latter result is not proved formally here for the sake of space, but follows straightforwardly from Anderson's Lemma (cf. Anderson, 1955).

3. SUBSAMPLING APPROXIMATION

We have seen before that the asymptotic null distribution of continuous functionals of R_n^1 depends in a complex way of the DGP and the specification under the null. Therefore, critical values for the test statistics can not be tabulated for general cases. In this section we overcome this problem with the assistance of the subsampling methodology. Resampling methods have been used extensively in the literature of quantile regression models, see, e.g., Hahn (1995), Horowitz (1998), Bilias, Chen and Ying (2000), Sakov and Bickel (2000) or He and Hu (2002). These articles consider *iid* sequences. When time series are involved the bootstrap approximation becomes more challenging. Subsampling is a powerful resampling scheme that allows an asymptotically valid inference under very general conditions on the DGP, see the monograph by Politis, Romano and Wolf (1999). Chernozhukov (2002) and Whang (2004) considered subsampling approximation for LQR models. In this section we apply the subsampling methodology to approximate the critical values of continuous functionals of R_n^1, thereby generalizing the aforementioned works to general nonlinear models. With an abuse of notation we write the test statistic as a function of the data $\{X_t = (Y_t, I_{t-1}')' : t = 0, \pm 1, \pm 2, ...\}$, $\Gamma(R_n^1) = \Gamma(R_n^1(X_1, ..., X_n))$. Let $G_n^\Gamma(w)$ be the test statistic cdf,

$$G_n^\Gamma(w) = P(\Gamma(R_n^1) \leq w).$$

Let $\Gamma(R_{b,i}^1) = \Gamma(R_b^1(X_i, ..., X_{i+b-1}))$ be the test statistic computed with the subsample $(X_i, ..., X_{i+b-1})$ of size b. We note that each subsample of size b (taken without replacement from the original data) is indeed a sample of size b from the true DGP. Hence, it is clear that one can approximate the sampling distribution $G_n^\Gamma(w)$ using the distribution of the values of $\Gamma(R_{b,i}^1)$ computed over the $n-b+1$ different subsamples of size b. That is, we approximate $G_n^\Gamma(w)$ by

$$G_{n,b}^\Gamma(w) = \frac{1}{n-b+1} \sum_{i=1}^{n-b+1} 1(\Gamma(R_{b,i}^1) \leq w), \qquad w \in [0, \infty).$$

Let $c_{n,1-\tau,b}^\Gamma$ be the $(1-\tau)$-th sample quantile of $G_{n,b}^\Gamma(w)$, i.e.,

$$c_{n,1-\tau,b}^\Gamma = \inf\{w : G_{n,b}^\Gamma(w) \geq 1 - \tau\}.$$

Thus, our subsampling tests reject the null hypothesis if $\Gamma(R_n^1) > c_{n,1-\tau,b}^\Gamma$. Let $c_{1-\tau}^\Gamma$ be the $(1-\tau)$-th quantile of $G_\infty^\Gamma(w) = P(\Gamma(R_\infty^1) \leq w)$. To justify theoretically this resampling approximation we need an additional assumption on the serial dependence of the DGP. Define the α-mixing coefficients as

$$\alpha(m) = \sup_{n \in \mathbb{Z}} \sup_{B \in \mathcal{F}_n, A \in \mathcal{P}_{n+m}} |P(A \cap B) - P(A)P(B)|, \; m \geq 1$$

where the σ-fields $\mathcal{F}_n$ and $\mathcal{P}_n$ are $\mathcal{F}_n := \sigma(X_t, t \leq n)$ and $\mathcal{P}_n := \sigma(X_t, t \geq n)$, respectively, with $X_t = (Y_t, Z_{t+1}')'$.

Assumption A6: $\{X_t = (Y_t, Z'_{t+1})' : t = 0, \pm 1, \pm 2, ...\}$ is a strictly stationary strong mixing process with α-mixing coefficients satisfying

$$\sum_{m=1}^{n} \alpha(m) = o(n).$$

The mixing assumption in A6 is sufficient but not necessary for the validity of the subsampling, see Politis, Romano and Wolf (1999). This subsampling procedure allows us to approximate the asymptotic critical values of the tests based on $\Gamma(R^1_{n,w})$. The next result justifies theoretically the subsampling approximation.

THEOREM 5: *Assume Assumptions A1-A6 and that $b/n \to 0$ and $b \to \infty$ as $n \to \infty$. Then,*

(i) Under the null hypothesis H_0,

$$c^{\Gamma}_{n,1-\tau,b} \xrightarrow{P} c^{\Gamma}_{1-\tau}.$$

and

$$P(\Gamma(R^1_n) > c^{\Gamma}_{n,1-\tau,b}) \longrightarrow \tau.$$

(ii) Under any fixed alternative hypothesis

$$P(\Gamma(R^1_n) > c^{\Gamma}_{n,1-\tau,b}) \longrightarrow 1.$$

(iii) Under the local alternatives (8),

$$P(\Gamma(R^1_n) > c^{\Gamma}_{n,1-\tau,b}) \longrightarrow P(\Gamma(R^1_\infty + D_a) > c^{\Gamma}_{1-\tau}).$$

Theorem 5 implies that the proposed subsampling tests have a correct asymptotic level, are consistent and are able to detect alternatives tending to the null at the parametric rate $n^{-1/2}$. An appealing property of our subsampling tests is that they do not need estimation of the nonparametric (conditional) sparsity function, which results in a substantial simplification of the tests. In practice, the empirical size and power of the tests depend on the choice of the parameter b. For this choice the reader is referred to Politis, Romano and Wolf (1999) or Sakov and Bickel (2000). In the present article, we follow the suggestion of Sakov and Bickel (2000) and we chose $b = \lfloor kn^{2/5} \rfloor$, where $\lfloor \cdot \rfloor$ denotes the integer part, which yields the optimal minimax accuracy under certain conditions. Section 5 below shows that this resampling procedure provides good approximations in finite samples for a variety of values for k.

4. DISTRIBUTION-FREE TESTS FOR LOCATION-SCALE MODELS

In this section we explore a rather different approach to study specification tests for the most popular class of models in the econometrics and statistical literature, the *location-scale* models. The main contribution of this section is to develop ADF specification tests for quantile regressions of such models based on certain weighted residual empirical processes.

Location-scale models are defined as

$$Y_t = f(I_{t-1}, \beta_0) + h(I_{t-1}, \beta_0)u_t(\beta_0), \tag{9}$$

where $\{u_t(\beta_0) \equiv u_t\}$ is a sequence of *iid* standardized errors, with u_t independent of I_{t-1}, and β_0 is an unknown finite-dimensional parameter in $\mathbb{R}^{p-1}$, $p > 1$. For these models, the associated conditional quantile is

$$m(I_{t-1}, \theta_0(\alpha)) = f(I_{t-1}, \beta_0) + h(I_{t-1}, \beta_0)F_u^{-1}(\alpha), \tag{10}$$

where F_u^{-1} is the quantile function of u_t, so the corresponding θ_0 is $\theta_0(\alpha) = (\beta_0', F_u^{-1}(\alpha))'$. Let f_u be the density of F_u.

Within this context, H_0 is equivalent to

$$E[1(u_t(\beta_0) \leq F_u^{-1}(\alpha)) - \alpha] = 0 \text{ for some } \theta_0(\alpha) = (\beta_0, F_u^{-1}(\alpha))' \in \Theta \subset \mathbb{R}^p, \forall \alpha \in \mathcal{T}.$$

Then it is natural to based a test on the weighted standardized residual empirical process

$$K_{n,w}^1(\alpha) := \frac{1}{\sqrt{n}} \sum_{t=1}^{n} w(I_{t-1}, \beta_n)\{1(u_t(\beta_n) \leq F_{u,n}^{-1}(\alpha)) - \alpha\},$$

where $w(I_{t-1}, \beta_0)$ is a real-valued measurable transformation of I_{t-1} that will be specified later on and depends on β_0, $u_t(\beta_n)$ are standardized residuals obtained from (10) using a $\sqrt{n}$-consistent estimator β_n, say, and $F_{u,n}^{-1}(\alpha)$ estimates $F_u^{-1}(\alpha)$. We can consider $\theta_{KB,n}(\alpha)$ for $\theta_0(\alpha) = (\beta_0, F_u^{-1}(\alpha))'$, but any other estimator satisfying A3 is possible, e.g. the Quasi-Maximum Likelihood Estimator (QMLE) for β_0 and the empirical quantile of residuals for F_u^{-1}.

Under H_0, $K_{n,w}^1$ is asymptotically centered, but under the alternative it is not asymptotically centered anymore, suggesting to base omnibus tests on suitable functionals of $K_{n,w}^1$. We choose the weights $w(I_{t-1}, \theta_n)$ and construct functionals in a simple way such that ADF tests are obtained, avoiding either subsampling approximations or complicated martingale transforms.

In the proof of Theorem 6 below we obtain, under the null H_0 and regularity conditions, the asymptotic uniform (in $\alpha \in \mathcal{T}$) expansion,

$$\begin{aligned} K_{n,w}^1(\alpha) &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} w(I_{t-1}, \beta_0)\{1(u_t(\beta_0) \leq F_u^{-1}(\alpha)) - \alpha\} \\ &\quad + \sqrt{n}(F_{u,n}^{-1}(\alpha) - F_u^{-1}(\alpha)) f_u(F_u^{-1}(\alpha)) E[w(I_{t-1}, \beta_0)] \\ &\quad + \sqrt{n}(\beta_n - \beta_0)' b(\alpha, w, \beta_0) + o_P(1), \end{aligned} \tag{11}$$

where

$$b(\alpha, w, \beta_0) := f_u(F_u^{-1}(\alpha)) E\left[w(I_{t-1}, \beta_0) a_{1,t}(\beta_0)\right] + E\left[w(I_{t-1}, \beta_0) a_{2,t}(\beta_0)\right] f_u(F_u^{-1}(\alpha)) F_u^{-1}(\alpha), \tag{12}$$

$$a_{1,t}(\beta) = \dot{f}_t(\beta)/h(I_{t-1}, \beta), \qquad a_{2,t}(\beta) = \dot{h}_t(\beta)/h(I_{t-1}, \beta),$$

with $\dot{f}_t(\beta_0) = \partial f(I_{t-1}, \beta_0)/\partial\beta$ and $\dot{h}_t(\beta_0) = \partial h(I_{t-1}, \beta_0)/\partial\beta$.

Notice that

$$E\left[w(I_{t-1}, \beta_0)a_{i,t}(\beta_0)\right] = 0 \qquad i = 1, 2, \tag{13}$$

implies $b(\alpha, w, \beta_0) \equiv 0$, and if in addition $w(I_{t-1}, \beta_0)$ has also zero mean, then a suitable standardization of $K^1_{n,w}(\alpha)$ is ADF.

To simplify notation write

$$X_t(\beta) := (1, a'_{1,t}(\beta), a'_{2,t}(\beta))' \qquad t = 1, ..., n.$$

To guarantee that (13) holds, we start with an initial $w(I_{t-1})$ and we shall take as $w(I_{t-1}, \beta_n)$ the residuals from the least squares regression (provided no exact collinearity exists, otherwise remove the necessary regressors),

$$w(I_{t-1}) = \gamma' X_t(\beta_n) + \xi_t \qquad t = 1, ..., n. \tag{14}$$

The initial w is up-to the econometrician and gives flexibility to direct the power of the tests against desired directions, see the end of this section.

The least squares estimator in (14) is

$$\widehat{\gamma}_n(\beta_n) = \left(\sum_{t=1}^{n} X_t(\beta_n)X'_t(\beta_n)\right)^{-1} \sum_{t=1}^{n} X_t(\beta_n)w(I_{t-1}).$$

The estimator $\widehat{\gamma}_n(\beta_n)$ estimates $\gamma \equiv \gamma(\beta_0) = \left(E[X_t(\beta_0)X'_t(\beta_0)]\right)^{-1} E[X_t(\beta_0)w(I_{t-1})]$ and the weight

$$w(I_{t-1}, \beta_0) = w(I_{t-1}) - \gamma'(\beta_0)X_t(\beta_0) \tag{15}$$

satisfies (13) and has zero mean, by construction. The function $\widehat{w}(I_{t-1}, \beta_n) = w(I_{t-1}) - \widehat{\gamma}'_n(\beta_n)X_t(\beta_n)$ estimates $w(I_{t-1}, \beta_0)$ in (15). Our final process is

$$K^1_{n,\widehat{w}}(\alpha) := \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \widehat{w}(I_{t-1}, \beta_n) 1(u_t(\beta_n) \leq F^{-1}_{u,n}(\alpha)).$$

To study the asymptotic behaviour of $K^1_{n,\widehat{w}}$ we need the following regularity conditions.

Assumption A7: Let Θ_{β_0} be a small convex neighborhood of β_0.

A7(a): The functions $f(I_{t-1}, \beta)$ and $h(I_{t-1}, \beta)$ are (a.s.) twice continuously differentiable in Θ_{β_0}. In addition,

$$E\left[\sup_{\beta \in \Theta_{\beta_0}} |X_t(\beta)|^2\right] < C,$$

and $I(\beta) = E\left[X_t(\beta)X'_t(\beta)\right]$ is positive definite on Θ_{β_0}.

A7(b): For a sufficiently small $\delta > 0$ and all sufficiently large n on,

$$E\left[\sup_{|\beta_1 - \beta_2| \leq \delta n^{-1/2}} |X_t(\beta_1) - X_t(\beta_2)|\right] \leq C\delta.$$

Assumption A8:

A8(a): $E\left[w^2(I_{t-1})\right] < C$.

A8(b): F_u is strictly increasing and has a Lebesgue density f_u that is uniformly bounded, i.e.,

$$\sup_{x\in\mathbb{R}} |f_u(x)| \leq C,$$

and equicontinuous: for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\sup_{|x-z|\leq\delta} |f_u(x) - f_u(z)| \leq \epsilon.$$

Assumption A7 is necessary to show that the estimation of $w(I_{t-1}, \beta_0)$ in (15) has not asymptotic effect on the limit process of $K^1_{n,\widehat{w}}$. Koul and Ling (2006) have shown that A7 is satisfied for most common examples in the literature, e.g. ARMA-GARCH models. The conditions in Assumption A8 are analogous to Assumption A1(d), having the same role. Set $\sigma^2 = E[w^2(I_0, \beta_0)]$.

Theorem 6: *Under the location-scale model (10), Assumptions A1(a), A3, and A7-A8*

$$K^1_{n,\widehat{w}}(\cdot) \Longrightarrow \sigma B(\cdot) \text{ in } \ell^\infty(\mathcal{T}),$$

where B is a standard Brownian Bridge on $[0,1]$.

An application of the CMT yields

$$CvM_{n,ls} := \int_{\mathcal{T}} \widehat{\sigma}^{-2} \left|K^1_{n,\widehat{w}}(\alpha)\right|^2 d\alpha \xrightarrow{d} \int_{\mathcal{T}} |B(\alpha)|^2 d\alpha, \tag{16}$$

and

$$KS_{n,ls} := \sup_{\alpha\in\mathcal{T}} \left|\widehat{\sigma}^{-1} K^1_{n,\widehat{w}}(\alpha)\right| \xrightarrow{d} \sup_{\alpha\in\mathcal{T}} |B(\alpha)|, \tag{17}$$

where $\widehat{\sigma}^2 = n^{-1}\sum_{t=1}^n \widehat{w}^2(I_{t-1}, \beta_n)$ estimates σ^2. The asymptotic critical values of the test statistics $CvM_{n,ls}$ and $KS_{n,ls}$ are distribution-free and can be easily tabulated, see Section 5.

A natural candidate for estimating F_u is the empirical cdf of residuals $F_{u,n}$. In such a case the test statistics $CvM_{n,ls}$ and $KS_{n,ls}$ can be easily computed, as the process $K^1_{n,\widehat{w}}(\alpha)$ takes at most $n - 2i_\epsilon + 1$ values, where $i_\epsilon = \inf\{j : 1 \leq j \leq n,\, j/n \geq \epsilon\}$. Similarly, if $F_{u,n}$ is a continuous estimator of F_u, $CvM_{n,ls}$ can be easily computed and no numerical integration is necessary.

The choice of the initial w allows us to construct omnibus ADF tests with power against desired alternatives. To illustrate this point, consider the following local alternatives within the model (9):

$$H_{1n} : E[Y_t \mid I_{t-1} = x] = f(x, \beta_0) + n^{-1/2} s_m(x). \tag{18}$$

Under these local alternatives the expansion in (11) typically is a sum of centered *iid* random variables plus a shift, see Behnen and Neuhaus (1975). The shift is in charge of local power. It can be shown that the shift (in absolute value) is

$$D_1(\alpha) = E\left[w(I_{t-1}, \beta_0) s_m(I_{t-1}) h^{-1}(I_{t-1}, \beta_0)\right] f_u(F_u^{-1}(\alpha)).$$

It is then clear that the optimal choice for $w(I_{t-1}, \beta_0)$ is the orthogonal projection of $s_m(I_{t-1})h^{-1}(I_{t-1}, \beta_0)$ on the orthocomplement of the span generated by $X_t(\beta_0)$.

As compared to other methods for obtaining ADF tests our tests are much simpler to compute. The weight function $\widehat{w}(I_{t-1}, \beta_n)$ can be estimated with any regression package and no nonparametric estimation of the (conditional) sparsity function is necessary. In contrast, martingale transforms require nonparametric estimations of this function in the computation of the scores, which may result in inaccurate size performance in finite samples.

5. FINITE SAMPLE PERFORMANCE

We investigate in this section, by means of a Monte Carlo experiment, the finite sample performance of the proposed tests. Our interest in these simulations is in the comparison between the ADF tests and the subsampling-based tests. We describe our simulation setup.

The choice of $\Phi(\cdot)$ in (5) is up-to the practitioner and gives flexibility to direct the power against some preferred alternatives. Following Escanciano and Velasco (2006) and references therein, we choose $\Phi(\cdot)$ equal to the $d-$variate standard normal random vector. Thus, our CvM test boils down to

$$CvM_n = n^{-1} \sum_{t=1}^{n} \sum_{s=1}^{n} \left(\int_{\mathcal{T}} \Psi_\alpha(Y_t - m(I_{t-1}, \theta_n)) \Psi_\alpha(Y_s - m(I_{s-1}, \theta_n)) dW(\alpha) \right) \exp(-\frac{1}{2} \left| I_{t-1} - I_{s-1} \right|^2).$$

We consider as W a uniform discrete distribution over a grid of $\mathcal{T}$ in $m = 21$ equidistributed points from ϵ to $1-\epsilon$. Denote by $\mathcal{T}_m = \{\alpha_j\}_{j=1}^{m}$ the points in the grid, with $\epsilon = \alpha_1 < \cdots < \alpha_m = 1 - \epsilon$.

We compute $CvM_{n,ls}$ and $KS_{n,ls}$ as

$$CvM_{n,ls} = \frac{(1-2\epsilon)}{\widehat{\sigma}^2(m-1)} \sum_{j=1}^{m-1} \left| K^1_{n,\widehat{w}}(\alpha_j) \right|^2,$$

and

$$KS_{n,ls} = \max_{\alpha \in \mathcal{T}_m} \left| \widehat{\sigma}^{-1} K^1_{n,\widehat{w}}(\alpha) \right|,$$

with initial weights w given below.

The limit processes in (16) and (17) are functionals of the Brownian Bridge on [0,1]. To approximate the critical values of such functionals we carry out simulations based on the so-called

Kac-Siegert expansion of the Brownian Bridge, i.e.

$$B(\cdot) = \sum_{j=1}^{\infty} \lambda_j^{1/2} \epsilon_j \psi_j(\cdot), \tag{19}$$

where

$$\lambda_j = \frac{1}{(j\pi)^2}, \qquad \psi_j(t) = \sqrt{2}\sin(j\pi t),\ t \in [0,1],\ j = 1, 2, ...,$$

and $\{\epsilon_j\}_{j=1}^{\infty}$ are *iid* $N(0,1)$ r.v's. We approximate the series in (19) using the first $r = 1,000$ summands of the series. Tables I and II report the approximated asymptotic critical values for $CvM_{n,ls}$ and $KS_{n,ls}$ for different values of m and based on 100,000 replications. As expected, the approximated critical values for $KS_{n,ls}$ are more sensitive to the choice of m than those of $CvM_{n,ls}$, especially for small values of m. Notice also that for large values of m, the asymptotic critical values for $KS_{n,ls}$ are very similar to those of the standard KS test of the Brownian Bridge on [0,1].

Please, insert Table I and Table II about here.

For the simulations, we examined two data generating processes that have been previously considered in Zheng (1998) and Whang (2004):

$$DGP1 : Y_t = X_{1t} + X_{2t} + c_1\sigma_t^{3/2} + u_{1t}, \qquad t = 1, \ldots, n,$$

where $\sigma_t = X_{1t}^2 + X_{2t}^2 + X_{1t}X_{2t}$ and X_{1t}, X_{2t} and $u_{1t} \sim iid\ N(0,1)$, mutually independent. The null hypothesis corresponds to the location model with $c_1 = 0$, so the null quantile model is a LQR model

$$m(I_{t-1}, \theta(\alpha)) = Z_t'\theta_0(\alpha), \qquad \alpha \in \mathcal{T},$$

with $Z_t = (1, X_{1t}, X_{2t})'$ and $\theta_0(\alpha) = (\phi^{-1}(\alpha), 1, 1)'$, with $\phi^{-1}(\alpha)$ the quantile function of the standard normal r.v.

The second design is a time series model:

$$DGP2 : Y_t = 0.6Y_{t-1} + X_t + c_2X_t^2 + u_{2t}, \qquad t = 1, \ldots, n,$$

where $X_t = 0.5X_{t-1} + \varepsilon_t$ with both u_{2t} and ε_t are sampled independently from $N(0,1)$ and $Y_0 = X_0 = 0$. Here, the null model corresponds to $c_2 = 0$. Under H_0, a LQR model holds with $I_{t-1} = (1, Y_{t-1}, X_t)'$, and $\theta_0(\alpha) = (\phi^{-1}(\alpha), 0.6, 1)'$.

We consider two sample sizes $n = 100$ and $n = 300$ and a quantile interval $[0.1, 0.9]$. As the number of subsamples, we follow the suggestion of Sakov and Bickel (2000) and we chose $b = \lfloor kn^{2/5} \rfloor$, with k from 9 to 11 for DGP1 and from 3 to 5 for DGP2, that yields for DGP1 (DGP2), $b = 54, 60$ and 66 $(18, 24,$ and $30)$ for $n = 100$ and $b = 81, 90$ and 99 $(27, 36$ and $45)$ for $n = 300$. We set the number of Monte Carlo repetitions to 1,000. The parameter $\theta_0(\alpha)$ is estimated by the QRE of Koenker and

Bassett (1978). In all experiments, the nominal probability of rejecting a correct null hypothesis is 0.05. The results with other nominal values are similar. To compute $CvM_{n,ls}$ and $KS_{n,ls}$, we choose $w_1(I_{t-1}) = X_{1t}^2$ and $w_2(I_{t-1}) = \sigma_t$ for DGP1 and $w_1(I_{t-1}) = |Y_{t-1}X_t|$ and $w_2(I_{t-1}) = X_t^2$ for DGP2. In tables we denote by $CvM_{n,i}$ and $KS_{n,i}$ the test statistics based on $w_i(I_{t-1})$, $i = 1, 2$.

Table III provides the rejection probabilities of the tests for DGP1. When $c_1 = 0$, the results show that the size performance of the subsampling-based test is good for all the subsample sizes considered and that the approximated asymptotic critical values lead to accurate empirical sizes for the ADF tests. We observe that to achieve appropriate empirical sizes the choice of b for the DGP1 should be larger than for the DGP2. When $c_1 \neq 0$, the results show the power performance of the tests. The rejection probabilities increase as n increases, as expected, showing that the tests are consistent against these fixed alternatives. For DGP1 the ADF tests outperform the subsampling-based test, with $CvM_{n,2}$ and $KS_{n,2}$ having the best empirical power, which is consistent with our local-power analysis. The latter conclusion was expected because the ADF tests take into account the location-structure of the model, and use of this information should produce better power properties. For the subsampling-based tests the power does not depend substantially on the choice of b. Table IV gives the corresponding results for DGP2 with similar conclusions to those under DGP1.

Unreported simulations using the indicator weight function $1(I_{t-1} \leq x)$, instead of $\exp(ix'I_{t-1})$, confirm that exponential-based tests have more power than indicator-based tests for these alternatives. In fact, this was our motivation for the use of the exponential weight in the CvM test.

This small simulation study suggests that even with relative small sample sizes the subsampling and ADF tests proposed in this article exhibit fairly good size accuracy and power.

Please, insert Table III and Table IV about here.

6. APPLICATIONS

In this section, we apply the new proposed tests for testing the correct specification of some well-known quantile models considered in the literature. More concretely, we examine two applications: first, we consider the Sharpe Style Analysis of the Magellan Fund studied in Kim and White (2003), see also Basset and Chen (2001), and second, the Pennsylvania Reemployment Bonus Experiments analyzed in Koenker and Xiao (2002). In both applications LQR models have been considered for a range of quantiles in $\mathcal{T} = [\epsilon, 1 - \epsilon]$ for a given $\epsilon \in (0, 0.5)$.

6.1 Application to Sharpe Style Analysis

Since Sharpe's (1988, 1992) seminal work, the Sharpe style regression has become a popular tool to analyze the style of an investment fund. The Sharpe style regression is carried out by regressing fund returns on various factors mimicking relevant indices. By analyzing the regression coefficients of the factors, one can understand the style of a fund manager. Bassett and Chen (2001) have proposed using the quantile regression method to analyze the style of a fund manager over the entire conditional distribution. These authors consider a linear specification

$$R_t = \theta_{00}(\alpha) + \theta_{10}(\alpha)Z_t^{LG} + \theta_{20}(\alpha)Z_t^{LV} + \theta_{30}(\alpha)Z_t^{SG} + \theta_{40}(\alpha)Z_t^{SV} + \varepsilon_{t,\alpha}, \tag{20}$$

where $\{R_t\}$ are the returns of the Fidelity Magellan fund, the factors are the Russell indices classified as:

	Large (L)	Small (S)
Growth (G)	Russell 1000 Growth (Z_t^{LG})	Russell 2000 Growth (Z_t^{SG})
Value (V)	Russell 1000 Value (Z_t^{LV})	Russell 2000 Value (Z_t^{SV})

The sample we consider is from January 1979 to December 1997, as in Kim and White (2003), with a total of 228 monthly observations. Details about the estimation and other related issues for this data set can be found in Basset and Chen (2001) for a shorter period and in Kim and White (2003) for the period considered here. In this section we are interested in testing the correct specification of the LQR model in (20) and to test if a pure location model is appropriate for this data set. Kim and White (2003) did not find evidence against the LQR specification. We consider $\epsilon = 0.1$ and $m = 9$, i.e., $\alpha = 0.1, 0.2, ..., 0.9$. For the subsampling we choose b between 90 and 100. We do not find evidence against the linear specification with the CvM subsampling-based test. The smallest empirical p-value for CvM subsampling-based test is 0.4388. As for the tests for a location model, we have considered as the initial weight w in the ADF tests the product of all possible combinations among pairs of regressors, i.e., $w(Z_t) = Z_t^i Z_t^j$, $i, j = LG$, LV, SG and SV, measuring all the interactions among regressors. None of the ADF tests find evidence against the pure location model. The maximum value for the test statistics are 0.7767 and 0.1349 for the CvM and KS tests, respectively, and they are attained at $w(I_{t-1}) = \left(Z_t^{LV}\right)^2$. These correspond approximately to p-values of 0.30 and 0.40, respectively. Therefore, our application suggests that the LQR model is correctly specified, and moreover, a pure location model seems to be a good model for this data set.

6.2 The Pennsylvania Reemployment Bonus Experiments

In this section we shall reanalyze the Pennsylvania reemployment bonus experiment conducted by the U.S. Department of Labor in the 1980's in order to test the incentive effects of alternative compensation schemes for the unemployment insurance (UI). There have been a large significant empirical and theoretical literature focusing on this data set and similar experiments, see Koenker and Xiao (2002) or Chernozhukov (2002) and references therein. In these controlled experiments, UI claimants were randomly offered a cash bonus if they find a job within some prespecified period of time and if the job was retained for a specified duration. The objective of these experiments was to evaluate the impact of such a scheme on the unemployment duration.

As in the aforementioned studies, we focus here on the compensation schedule that includes a lump-sum payment of a six times the weekly unemployment benefit for claimants establishing the reemployment within 12 weeks (in addition to the usual weekly benefits). The definition of unemployment spell includes one waiting week, with the maximum of interrupted full weekly benefits of 27. The number of observations is 6384.

Koenker and Xiao (2002) fitted to this data set the linear quantile specification

$$Y_t = \theta_{00}(\alpha) + \theta_{01}(\alpha) D_t + \theta_{02}'(\alpha) X_t + \varepsilon_{t,\alpha}, \tag{21}$$

where Y_t is the log of the duration of unemployment, i.e., $Y_t = \log(T_t)$, D_t is the indicator of the bonus offer, and X_t is a set of socio-demographic characteristics (age, gender, number of dependents, location within the state, existence of recall expectations, and type of occupation). See Koenker and Xiao (2002) for a detailed analysis of this data set.

In Koenker and Xiao (2002) the interest was mainly in testing for restrictions on the parameter $\theta_0(\alpha)$ in the LQR model (21), e.g. testing for a pure location model, testing if the treatment effect is constant across the range of quantiles of interest and whether the treatment was unambiguously beneficial. Here in the present article we are concerned with testing if the LQR model is correctly specified.

We set $\epsilon = 0.15$ and $m = 15$ and compute our CvM subsampling test for this data set taking $b = 3000$ (see Chernozhukov (2002) for motivation on this choice). We have obtained an empirical $p-$value of 0 with the subsampling test, and hence, CvM_n strongly rejects the LQR specification. Other values of ϵ, m and b yield the same conclusion. For the pure location model, our results based on the ADF tests coincide with those obtained by Koenker and Xiao (2002) and Chernozhukov (2002), rejecting the pure location model. The KS test using $w(Z_t) = D_t B_t$, where B_t is a dummy variable which is 1 if the individual is black and 0 otherwise, rejects the location model at 10% and 5% with a value of 1.353, but the CvM test does not find evidence against the location model for this choice of w. Other choices of w lead to stronger rejections by ADF tests. For instance, the choice $w(Z_t) = F_t d_t$, where F_t is a dummy variable for gender (1 if female, 0 otherwise) and d_t is the number of dependents, lead to rejections with both, the KS and CvM tests, with respective values 1.523 and 0.566, and confirming the need of an interaction term between these two variables, as expected given the nature of the experiment.

Summarizing, we find evidence against the LQR model with our subsampling-based test and against the pure location model with the ADF tests. Notice that the asymptotic properties (e.g. consistency) of the estimator of $\theta_0(\alpha)$ in (21) are not necessarily affected by the misspecification of the LQR but are at least questionable. More concretely, even if the LQR model is misspecified it is still possible that $\theta_n(\alpha)$ in (21) estimates consistently $\theta_0(\alpha)$ defined by the moment conditions

$$E[\Psi_\alpha(Y_t - \theta_0'(\alpha)Z_t)Z_t] = 0, \qquad \forall \alpha \in \mathcal{T}, \tag{22}$$

see Kim and White (2003). But more importantly, it is possible that under misspecification of (21) still the condition

$$q_\alpha(D_t = 1, X_t) - q_\alpha(D_t = 0, X_t) = \theta_{01}(\alpha) \text{ a.s for all } \alpha \in \mathcal{T},$$

holds, which is the object of interest in this experiment. If the concern is not in testing the validity of the LQR model against *all* alternatives, but in testing the LQR model against those alternatives where $q_\alpha(D_t = 1, X_t) - q_\alpha(D_t = 0, X_t)$ and $\theta_{01}(\alpha)$ differ, more efficient tests taking into account that information are possible, see Escanciano and Song (2006) for a related problem in a different semiparametric testing setup. The development of such efficient tests in the present context is an interesting problem that deserves further attention and is a direction of future research.

APPENDIX. PROOFS

First, we shall state a weak convergence theorem which is an extension of Theorem A1 in Delgado and Escanciano (2006) and that is of independent interest. Let for each $n \geq 1$, $I'_{n,0}, ..., I'_{n,n-1}$, be an array of random vectors in $\mathbb{R}^p$, $p \in \mathbb{N}$, and $Y_{n,1}, ..., Y_{n,n}$, be an array of real random variables (r.v.'s). Denote by $(\Omega_n, \mathcal{A}_n, P_n)$, $n \geq 1$, the probability space in which all the r.v.'s $\{Y_{n,t}, I'_{n,t-1}\}_{t=1}^n$ are defined. Let $\mathcal{F}_{n,t}$, $0 \leq t \leq n$, be a double array of sub σ-fields of $\mathcal{A}_n$ such that $\mathcal{F}_{n,t-1} \subset \mathcal{F}_{n,t}$, $t = 1, ..., n$ and such that for each $n \geq 1$ and each $\gamma \in \mathcal{H}$,

$$E[w(Y_{n,t}, I_{n,t-1}, \gamma) \mid \mathcal{F}_{n,t-1}] = 0 \text{ a.s.} \qquad 1 \leq t \leq n, \ \forall n \geq 1. \tag{23}$$

Moreover, we shall assume that $\{w(Y_{n,t}, I_{n,t-1}, \gamma), \mathcal{F}_{n,t}, 0 \leq t \leq n\}$ is a square-integrable martingale difference sequence for each $\gamma \in \mathcal{H}$, that is, (23) holds, $Ew^2(Y_{n,t}, I_{n,t-1}, \gamma) < \infty$ and $w(Y_{n,t}, I_{n,t-1}, \gamma)$ is $\mathcal{F}_{n,t}$-measurable for each $\gamma \in \mathcal{H}$ and $\forall t, 1 \leq t \leq n, \forall n \in \mathbb{N}$. The following result gives sufficient conditions for the weak convergence of the empirical process

$$\alpha_{n,w}(\gamma) = n^{-1/2} \sum_{t=1}^{n} w(Y_{n,t}, I_{n,t-1}, \gamma) \qquad \gamma \in \mathcal{H}.$$

Under mild conditions the empirical process $\alpha_{n,w}$ can be viewed as a mapping from Ω_n to $\ell^\infty(\mathcal{H})$, the space of all complex-valued functions that are uniformly bounded on $\mathcal{H}$, with $\mathcal{H}$ a generic metric space. The weak convergence theorem that we present here is funded on results by Levental (1989), Bae and Levental (1995) and Nishiyama (2000). In Theorem A1 in Delgado and Escanciano (2006) $\mathcal{H}$ was finite-dimensional, but here we allow for an infinite-dimensional $\mathcal{H}$. The proof of theorem does not change by this possibility, however.

An important role in the weak convergence theorem is played by the conditional quadratic variation (CV) of the empirical process $\alpha_{n,w}$ on a finite partition $\mathcal{B} = \{H_k; 1 \leq k \leq N\}$ of $\mathcal{H}$, which is defined as

$$CV_{n,w}(\mathcal{B}) = \max_{1 \leq k \leq N} n^{-1} \sum_{t=1}^{n} E\left[\sup_{\gamma_1, \gamma_2 \in H_k} |w(Y_{n,t}, I_{n,t-1}, \gamma_1) - w(Y_{n,t}, I_{n,t-1}, \gamma_2)|^2 \mid \mathcal{F}_{n,t-1} \right]. \tag{24}$$

Then, for the weak convergence theorem we need the following assumptions.

W1: For each $n \geq 1$, $\{(Y_{n,t}, I_{n,t-1})' : 1 \leq t \leq n\}$ is a strictly stationary and ergodic process. The sequence $\{w(Y_{n,t}, I_{n,t-1}, \gamma), \mathcal{F}_{n,t}, 1 \leq t \leq n\}$ is a square-integrable martingale difference sequence for each $\gamma \in \mathcal{H}$. Also, there exists a function $C_w(\gamma_1, \gamma_2)$ on $\mathcal{H} \times \mathcal{H}$ to $\mathbb{R}$ such that uniformly in $(\gamma_1, \gamma_2) \in \mathcal{H} \times \mathcal{H}$

$$n^{-1} \sum_{t=1}^{n} w(Y_{n,t}, I_{n,t-1}, \gamma_1) w^c(Y_{n,t}, I_{n,t-1}, \gamma_2) = C_w(\gamma_1, \gamma_2) + o_{P_n}(1).$$

W2: The family $w(Y_{n,t}, I_{n,t-1}, \gamma)$ is such that $\alpha_{n,w}$ is a mapping from Ω_n to $\ell^{\infty}(\mathcal{H})$ and for every $\delta > 0$ there exists a finite partition $\mathcal{B}_\delta = \{H_k; 1 \leq k \leq N_\delta\}$ of $\mathcal{H}$, with N_δ being the elements of such partition, such that

$$\int_0^\infty \sqrt{\log(N_\delta)} d\delta < \infty \tag{25}$$

and

$$\sup_{\delta \in (0,1) \cap \mathbb{Q}} \frac{CV_{n,w}(\mathcal{B}_\delta)}{\delta^2} = O_{P_n}(1). \tag{26}$$

Let $\alpha_{\infty,w}(\cdot)$ be a Gaussian process with zero mean and covariance function given by $C_w(\gamma_1, \gamma_2)$. We are now in position to state the following

Theorem A1: *If Assumptions W1 and W2 hold, then it follows that*

$$\alpha_{n,w} \Longrightarrow \alpha_{\infty,w} \text{ in } \ell^{\infty}(\mathcal{H}).$$

Proof of Theorem A1: Theorem A1 in Delgado and Escanciano (2006).

Corollary A1: *Assuming that W1 holds for* $w(Y_{n,t}, I_{n,t-1}, v) = \Psi_\alpha(Y_{n,t} - m(I_{n,t-1}, \theta_0(\alpha))) \exp(ix' I_{n,t-1})$, $v = (x', \alpha)' \in \Pi$, *A1(b) and that*

$$n^{-1} \sum_{t=1}^{n} |I_{n,t-1}|^2 = O_{P_n}(1),$$

then the weak convergence of Theorem A1 holds.

Proof of Corollary A1: We shall apply Theorem A1. Let us define the metric

$$d(v_1, v_2) := \sqrt{|\alpha_1 - \alpha_2| + |x_1 - x_2|^2}, \qquad v_1, v_2 \in \Pi.$$

Then, we define an δ-bracket as an interval $[v_1, v_2]$ such that $v_1 \leq v_2$ and $d(v_1, v_2) \leq \delta$. The bracketing number $N(\delta, \Pi, d)$ is the minimum number of δ-brackets needed to cover Π. Then, it is easy to show that

$$\int_0^\infty \sqrt{\log(N(\delta, \Pi, d))} d\delta < \infty$$

holds. It remains to show that (26) holds. Consider a partition $\mathcal{B}_\delta = \{H_k; 1 \leq k \leq N(\delta, \Pi, d) \equiv N_\delta\}$ of Π in δ-brackets $H_k = [\underline{v}_k, \overline{v}_k]$, with $\underline{v}_k = (\underline{x}_k', \underline{\alpha}_k)'$ and $\overline{v}_k = (\overline{x}_k', \overline{\alpha}_k)'$, $\underline{x}_k \leq \overline{x}_k$ and $\underline{\alpha}_k \leq \overline{\alpha}_k$. Define

$\varepsilon_{n,t}(\alpha) = Y_{n,t} - m(I_{n,t-1}, \theta_0(\alpha))$. Then, by simple algebra and the monotonicity of $1(\varepsilon_{n,t}(\alpha) \leq 0)$ due to A1(b), $CV_{n,w}(\mathcal{B}_\delta)$ in (24) is bounded by

$$
\begin{aligned}
& 2 \max_{1\leq k\leq N_\epsilon} n^{-1}\sum_{t=1}^{n} E\left[\sup_{v_1,v_2\in H_k} |1(\varepsilon_{n,t}(\alpha_1) \leq 0) - \alpha_1 - 1(\varepsilon_{n,t}(\alpha_2) \leq 0) + \alpha_2|^2 \mid \mathcal{F}_{n,t-1}\right] \\
& +2 \max_{1\leq k\leq N_\epsilon} n^{-1}\sum_{t=1}^{n} \left[\sup_{v_1,v_2\in H_k} |\exp(ix_1' I_{n,t-1}) - \exp(ix_2' I_{n,t-1})|^2\right] \\
\leq \; & C \max_{1\leq k\leq N_\epsilon} \left\{ |\underline{\alpha}_k - \overline{\alpha}_k| + |\underline{x}_k - \overline{x}_k|^2 n^{-1}\sum_{t=1}^{n} |I_{n,t-1}|^2 \right\}.
\end{aligned}
$$

Hence, (26) holds for the partition $\mathcal{B}_\delta$. Therefore, W2 of Theorem A1 holds and the corollary is proved. □

Proof of Theorem 1. Follows from Corollary A1. □

Theorem A2. *Assume Assumptions A1(c-d), A2, A3(a), and that there exists a* $\theta_1 \in \mathcal{B}$ *such that* $\|\theta_n - \theta_1\|_{\mathcal{B}} = o_P(1)$. *Then, uniformly in* $(x', \alpha)' \in \Pi$,

$$
\begin{aligned}
R_n^1(x,\alpha) \;=\; & \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \{\Psi_\alpha(e_t(\theta_1)) - E[\Psi_\alpha(e_t(\theta_1)) \mid \mathcal{F}_{t-1}]\} \exp(ix' I_{t-1}) \qquad (27) \\
& +\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \{E[\Psi_\alpha(e_t(\theta)) \mid \mathcal{F}_{t-1}]_{\theta=\theta_n} - E[\Psi_\alpha(e_t(\theta_1)) \mid \mathcal{F}_{t-1}]\} \exp(ix' I_{t-1}) \\
& +\frac{1}{\sqrt{n}}\sum_{t=1}^{n} E[\Psi_\alpha(e_t(\theta_1)) \mid \mathcal{F}_{t-1}] \exp(ix' I_{t-1}) - E\left[E[\Psi_\alpha(e_t(\theta_1)) \mid \mathcal{F}_{t-1}] \exp(ix' I_{t-1})\right] \\
& +\sqrt{n} E\left[E[\Psi_\alpha(e_t(\theta_1)) \mid \mathcal{F}_{t-1}] \exp(ix' I_{t-1})\right] + o_P(1).
\end{aligned}
$$

Proof of Theorem A2: Write $w_{t-1}(v,\theta) := \{\Psi_\alpha(e_t(\theta)) - E[\Psi_\alpha(e_t(\theta)) \mid \mathcal{F}_{t-1}]\} \exp(ix' I_{t-1})$. First we shall show that the process

$$
S_n(v,\theta) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_{t-1}(v,\theta)
$$

is asymptotically tight with respect to $(v,\theta) \in \mathcal{W} = \Pi \times \mathcal{B}$.

Let us define the class $\mathcal{K} = \{w_\cdot(v,\theta) : (v,\theta) \in \mathcal{W}\}$. Denote $X_{t-1,\infty} = (I_{t-1}, I_{t-2}, ...)'$. Let $\mathcal{B}_\delta = \{B_k; 1 \leq k \leq N_\delta \equiv N_{[]}(\delta, \mathcal{K}, \|\cdot\|_2\}$, with $B_k = [\underline{w}_k(Y_t, X_{t-1,\infty}), \overline{w}_k(Y_t, X_{t-1,\infty})]$, be a partition of $\mathcal{K}$ in δ-brackets with respect to $\|\cdot\|_2$, where $\|\cdot\|_2$ denotes the L_2 norm of random variables, i.e., $\|X\|_2 = \left(E[X^2]\right)^{1/2}$.

Conditions A1(c-d) and A2 imply that for a sufficiently small $\delta > 0$,

$$\left\| \sup_{\substack{(v_2,\theta_2)\in\mathcal{A}:d(v_1,v_2)\leq\delta \\ \|\theta_1-\theta_2\|_{\mathcal{B}}\leq\delta}} |w_{t-1}(v_1,\theta_1) - w_{t-1}(v_2,\theta_2)| \right\|_2 \tag{28}$$

$$\begin{aligned} &\leq C \left\| \sup_{\substack{(v_2,\theta_2)\in\mathcal{A}:d(v_1,v_2)\leq\delta \\ \|\theta_1-\theta_2\|_{\mathcal{B}}\leq\delta}} |\Psi_{\alpha_1}(e_t(\theta_1)) - \Psi_{\alpha_2}(e_t(\theta_2))| \right\|_2 + C\delta \\ &\leq C\delta^{1/2} + C\delta \leq C\delta^{1/2}. \end{aligned}$$

Theorem 3 in Chen et al. (2003) and A3(a) yield that (25) holds for such partition. Therefore, by similar arguments as in Corollary A1, (26) follows, and condition W2 of Theorem A1 holds. The asymptotically tightness of $S_n(v,\theta)$ is then proved.

Now, write

$$\begin{aligned} R_n^1(\cdot) &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{\Psi_\alpha(e_t(\theta_1)) - E[\Psi_\alpha(e_t(\theta_1)) \mid \mathcal{F}_{t-1}]\} \exp(ix'I_{t-1}) \\ &\quad + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} E[\Psi_\alpha(e_t(\theta)) \mid \mathcal{F}_{t-1}]_{\theta=\theta_n} + o_P(1), \end{aligned}$$

and (27) follows. $\square$

Proof of Theorem 2: Under the null $\theta_1 = \theta_0$ and $E[\Psi_\alpha(e_t(\theta_0)) \mid \mathcal{F}_{t-1}] = 0$ a.s. From the expansion in (27), it follows that, uniformly in $v \in \Pi$,

$$\begin{aligned} R_n^1(\cdot) &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Psi_\alpha(e_t(\theta_0)) \exp(ix'I_{t-1}) \\ &\quad + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{E[\Psi_\alpha(e_t(\theta)) \mid \mathcal{F}_{t-1}]_{\theta=\theta_n} - E[\Psi_\alpha(e_t(\theta_0)) \mid \mathcal{F}_{t-1}]\} \exp(ix'I_{t-1}) + o_P(1) \\ &= R_n(\cdot) + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\{F_{I_{t-1}}(m(I_{t-1},\theta_n)) - F_{I_{t-1}}(m(I_{t-1},\theta_0))\right\} \exp(ix'I_{t-1}) + o_P(1). \end{aligned}$$

Now, from A1(d) and Koul and Stute (1999, pp. 228-229), uniformly in $v \in \Pi$,

$$\begin{aligned} &\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\{F_{I_{t-1}}(m(I_{t-1},\theta_n)) - F_{I_{t-1}}(m(I_{t-1},\theta_0))\right\} \exp(ix'I_{t-1}) \\ &= \sqrt{n}(\theta_n - \theta_0) \frac{1}{n} \sum_{t=1}^{n} g(I_{t-1},\theta_0) f_{I_{t-1}}(m(I_{t-1},\theta_0)) \exp(ix'I_{t-1}) + o_P(1). \end{aligned}$$

This together with Theorem 1, A2(c) and A3 proves the theorem. $\square$

Proof of Theorem 3: From Theorem A2 and the Ergodic Theorem

$$\sup_{v\in\Pi} \left| \frac{1}{n} \sum_{t=1}^{n} [\Psi_\alpha(e_t(\theta_n(\alpha))) \exp(ix'I_{t-1}) - E[\Psi_\alpha(e_t(\theta_1(\alpha))) \exp(ix'I_{t-1})] \right| = o_P(1). \tag{29}$$

Let $\mathcal{W} = \Pi \times \mathcal{B}$. Let $w = (x', \alpha, \theta'(\cdot))'$ be a general element of $\mathcal{W}$. The space $\mathcal{W}$ is endowed with the metric

$$\rho(w_1, w_2) = |x_1 - x_2| + |\alpha_1 - \alpha_2| + \sup_{\alpha \in \mathcal{T}} |\theta_1(\alpha) - \theta_2(\alpha)|,$$

where $w_1 = (x_1', \alpha_1, \theta_1'(\cdot))'$ and $w_2 = (x_2', \alpha_2, \theta_2'(\cdot))'$ belong to $\mathcal{W}$. Let $B(w, \delta)$ be the open ball of radius δ around w, i.e., $B(w, \delta) = \{w_1 \in \mathcal{W} : \rho(w_1, w) < \delta\}$. Note that A1-A3 yield that for each $w = (x', \alpha, \theta'(\cdot))' \in \mathcal{W}$ it holds that

$$\lim_{\delta \to 0} E\left[\sup_{w_1 \in B(w,\delta)} \left|\Psi_{\alpha_1}(e_t(\theta_1(\alpha_1))) \exp(ix_1' I_{t-1}) - \Psi_\alpha(e_t(\theta(\alpha))) \exp(ix' I_{t-1})\right|^2\right] = 0.$$

Therefore, $E[\Psi_\alpha(e_t(\theta_1(\alpha))) \exp(ix' I_{t-1})]$ is a continuous function of $v = (x', \alpha)'$. Therefore, under the alternative H_A we have that the function $E[\Psi_\cdot(e_t(\theta_1(\cdot)))1(I_{t-1} \leq \cdot)]$ is different from zero in a subset with positive Lebesgue measure on Π. $\square$

Proof of Theorem 4: The proof follows from Theorem A2 and Assumptions A4 and A5 jointly with A3' in a routine fashion, and then, it is omitted. $\square$

Proof of Theorem 5. The proof follows the same steps as Theorems 2, 3 and 4 of Whang (2004) and then, it is omitted. $\square$

Before proving Theorem 6 we need a useful Lemma. To emphasize the dependence of $X_t(\beta)$ on I_{t-1}, we write when it is convenient $X_t(\beta) \equiv X(I_{t-1}, \beta)$. Notice that in the context of location-scale models $\theta_0(\alpha) = (\theta_{01}(\alpha), \theta_{02}(\alpha)) = (\beta_0', F_u^{-1}(\alpha))'$. Write similarly, $\theta_1(\alpha) = (\theta_{11}, \theta_{12}(\alpha))$ and $\theta_2(\alpha) = (\theta_{21}, \theta_{22}(\alpha))$. Define the process

$$K_n(\beta, \alpha, \theta) := \frac{1}{\sqrt{n}} \sum_{t=1}^{n} X(I_{t-1}, \beta)\{1(Y_t \leq m(I_{t-1}, \theta(\alpha)) - \alpha\}$$

indexed by $(\beta, \alpha, \theta) \in \mathcal{C}_{n,K} \times \mathcal{T} \times \mathcal{B}$, where $\mathcal{C}_{n,K}$ is a shrinking neighborhood of β_0 such that for a sufficiently large $K > 0$,

$$\mathcal{C}_{n,K} = \left\{\beta \in \Theta_\beta : \sqrt{n}|\beta - \beta_0| < K\right\}.$$

Lemma A1: In the context of the location-scale model in (9). Under Assumption A7, and that F_u is strictly increasing the process $K_n(\beta, \alpha, \theta)$ is asymptotically tight with respect to $(\beta, \alpha, \theta) \in \mathcal{C}_{n,K} \times \mathcal{T} \times \mathcal{B}$.

Proof of Lemma A1: Let us define the class of functions $\mathcal{K}_1 = \{X(I_{t-1}, \beta)\{1(Y_t \leq m(I_{t-1}, \theta(\alpha)) - \alpha\} : (\beta, \alpha, \theta) \in \mathcal{C}_{n,K} \times \mathcal{T} \times \mathcal{B}\}$. Denote now $X_{t-1,\infty} = (I_{t-1}, I_{t-2}, ...)'$. Let $\mathcal{B}_\delta = \{B_k; 1 \leq k \leq N_\delta \equiv N_{[]}(\delta, \mathcal{K}_1, \|\cdot\|_2)\}$, with $B_k = [\underline{w}_k(Y_t, X_{t-1,\infty}), \overline{w}_k(Y_t, X_{t-1,\infty})]$, be a partition of $\mathcal{K}_1$ in δ-brackets with respect to $\|\cdot\|_2$. Write $w_{t-1}(\beta,\alpha,\theta) = X(I_{t-1},\beta)\{1(Y_t \leq m(I_{t-1},\theta(\alpha)) - \alpha\}$. Condition A7 and triangle's inequality yield

$$
\begin{aligned}
& E\left|\sup_{\substack{(\beta_2,\alpha_2,\theta_2)\in\mathcal{A}:|\alpha_1-\alpha_2|\leq\delta \\ |\beta_1-\beta_2|\leq\delta,\|\theta_1-\theta_2\|_{\mathcal{B}}\leq\delta}} |w_{t-1}(\beta_1,\alpha_1,\theta_1) - w_{t-1}(\beta_2,\alpha_2,\theta_2)|\right| \\
\leq\ & C\left(E\left|X(I_{t-1},\beta_1)\right|^2\right)^{1/2}\left(|F_u(\theta_{12}(\alpha_1)-\delta)) - F_u(\theta_{12}(\alpha_1)+\delta)|\right)^{1/2} + C\delta \\
& + E\left|\sup_{|\beta_1-\beta_2|\leq\delta} |X(I_{t-1},\beta_1) - X(I_{t-1},\beta_2)|\right| \leq C\delta^{1/2},
\end{aligned}
$$

for a sufficiently small $\delta > 0$. Theorem 3 in Chen et al. (2003) and A3(a) yield that (25) holds for such partition. Therefore, by similar arguments as in Corollary A1, (26) follows, and condition W2 of Theorem A1 holds. The asymptotically tightness of $K_n(\beta,\alpha,\theta)$ is then proved. □

Proof of Theorem 6: Write $q_t(\alpha,\theta_n) := \{1(u_t(\beta_n) \leq F_{u,n}^{-1}(\alpha)) - \alpha\}$ and

$$
\begin{aligned}
K_{n,\widehat{w}}^1(\alpha) &= \frac{1}{\sqrt{n}}\sum_{t=1}^n \{w(I_{t-1}) - \widehat{\gamma}_n'(\beta_n)X_t(\beta_n)\}q_t(\alpha,\theta_n) \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^n \{w(I_{t-1}) - \gamma'(\beta_0)X_t(\beta_n)\}q_t(\alpha,\theta_n) \\
&\quad -(\widehat{\gamma}_n'(\beta_n) - \gamma'(\beta_0))\frac{1}{\sqrt{n}}\sum_{t=1}^n X_t(\beta_n)q_t(\alpha,\theta_n) \\
&:= I_{1n}(\alpha) + I_{2n}(\alpha).
\end{aligned}
$$

From A7 and the uniform law of large numbers of Jennrich (1969),

$$
\left|\widehat{\gamma}_n'(\beta_n) - \gamma'(\beta_0)\right| = o_P(1),
$$

and from Lemma A1,

$$
\sup_{\alpha\in\mathcal{T}}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^n X_t(\beta_n)q_t(\alpha,\theta_n)\right| = O_P(1).
$$

Hence $\sup_{\alpha\in\mathcal{T}}|I_{2n}(\alpha)| = o_P(1)$.

As for $I_{1n}(\alpha)$, again by Lemma A1 and writing $w(I_{t-1},\beta_0) := w(I_{t-1}) - \gamma'(\beta_0)X_t(\beta_0)$,

$$
\sup_{\alpha\in\mathcal{T}}\left|I_{1n}(\alpha) - \frac{1}{\sqrt{n}}\sum_{t=1}^n w(I_{t-1},\beta_0)q_t(\alpha,\theta_n)\right| = o_P(1).
$$

Define

$$
d_{1tn} := \frac{f(I_{t-1},\beta_n) - f(I_{t-1},\beta_0)}{h(I_{t-1},\beta_0)} \qquad d_{2tn} := \frac{h(I_{t-1},\beta_n) - h(I_{t-1},\beta_0)}{h(I_{t-1},\beta_0)},
$$

and $i_{t,n}(\alpha) := F_{u,n}^{-1}(\alpha) + d_{1tn} + d_{2tn}F_{u,n}^{-1}(\alpha)$.

Now,

$$
\frac{1}{\sqrt{n}}\sum_{t=1}^n w(I_{t-1},\beta_0)q_t(\alpha,\theta_n) := A_{1n}(\alpha) + A_{2n}(\alpha) + A_{3n}(\alpha), \tag{30}
$$

where

$$A_{1n}(\alpha) := \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w(I_{t-1}, \beta_0) q_t(\alpha, \theta_0),$$

$$A_{2n}(\alpha) := \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w(I_{t-1}, \beta_0)\{q_t(\alpha, \theta_n) - q_t(\alpha, \theta_0) - F_u(i_{t,n}(\alpha)) + \alpha\}$$

and

$$A_{3n}(\alpha) := \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w(I_{t-1}, \beta_0)\{F_u(i_{t,n}(\alpha)) - \alpha\}.$$

By similar arguments to those of Lemma A1 it can be shown that $\sup_{\alpha \in \mathcal{T}} |A_{2n}(\alpha)| = o_P(1)$. Whereas, from the arguments of Koul and Stute (1999, pp. 228-229), it can be shown that, uniformly in $\alpha \in \mathcal{T}$,

$$\begin{aligned} A_{3n}(\alpha) &= \sqrt{n}(F_{u,n}^{-1}(\alpha) - F_u^{-1}(\alpha)) f_u(F_u^{-1}(\alpha)) E[w(I_{t-1}, \beta_0)] \\ &\quad + \sqrt{n}(\beta_n - \beta_0)' b(\alpha, w, \beta_0) + o_P(1). \end{aligned} \tag{31}$$

The theorem follows from (30), (31) and Lemma A1. □

REFERENCES

Anderson, T.W. (1955): "The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities", *Proceedings of the American Mathematical Society,* 6, 170-176.

Angrist, J. Chernozhukov, V. and Fernandez-Val, I. (2006): "Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure", forthcoming in *Econometrica.*

Bae, J. and Levental, S. (1995): "Uniform CLT for Markov chains and its invariance principle: a martingale approach", *Journal of Theoretical Probability*, 8, 549-570.

Bassett, G. and Chen, H.L. (2001): "Portfolio style: return-based attribution using regression quantiles", *Empirical Economics*, 26, 293-305.

Bassett, Jr., G., and Koenker, R. (1978): "Asymptotic theory of least absolute error regression," *Journal of the American Statistical Association*, 73, 618–622.

Behnen, K. and Neuhaus, G. (1975): "A central limit theorem under contiguous alternatives", *The Annals of Statistics,* 3, 1349-1353.

Bierens, H.J. (1982): "Consistent model specification tests", *Journal of Econometrics*, 20, 105-134.

Bierens, H., and Ginther, D.K. (2001): "Integrated conditional moment testing of quantile regression models", *Empirical Economics*, 26, 307-324.

Bierens, H.J. and Ploberger, W. (1997): "Asymptotic theory of integrated conditional moment test", *Econometrica*, 65, 1129-1151.

Bilias, Y., Chen, S. and Ying, Z. (2000): "Simple resampling methods for the censored regression quantiles", *Journal of Econometrics*, 99, 373–386.

Billingsley, P. (1961). The Lindeberg-Levy theorem for martingales. *Proceedings of the American Mathematical Society*, 12, 788-792.

Chen, X., Linton, O. and van Keilegom, I. (2003): "Estimation of semiparametric models when the criterion function is not smooth", *Econometrica*, 71, 1591-1608.

Chernozhukov, V. (2002): "Inference on the quantile regression process, an alternative," Working article 02-12 (February), MIT (www.ssrn.com).

Delgado, M.A. and Escanciano, J.C. (2006): "Nonparametric tests for conditional symmetry in dynamic models", forthcoming in *Journal of Econometrics.*

Dudley, R.M. (1999): *Uniform Central Limit Theorems.* Cambridge. University Press, Massachusetts.

Escanciano, J.C. and Velasco, C. (2006): "Generalized spectral tests for the martingale difference hypothesis", *Journal of Econometrics*, 134, 151-185.

Escanciano, J.C and Song, K. (2006): “Asymptotically Optimal Tests for Single-Index Models with a Focus on Average Partial Effects”, unpublished manuscript.

Gutenbrunner, C. and Jurečkova, J. (1992): “Regression rank scores and regression quantiles”, *The Annals of Statistics*, 20, 305–329.

Gutenbrunner, C., Jurečkova, J., Koenker, R. and Portnoy, S. (1993): “Test of linear hypotheses based on regression rank scores”, *Journal of Nonparametric Statistics*, 2, 307-331.

Hahn, J. (1995): “Bootstrapping Quantile Regression Estimators”, *Econometric Theory*, 11, 105-121.

He, X., and Hu, F. (2002): “Markov chain marginal bootstrap”, *Journal of the American Statistical Association*, 97, 783–795

He, X. and Zhu, L.X (2003): “A Lack-of-Fit test for Quantile Regression”, *Journal of the American Statistical Association*, 98, 1013-1022.

Horowitz, J.L. (1998): “Bootstrap methods for median regression models”, *Econometrica*, 66, 1327-1351.

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.

Inoue, A. (1999): “A conditional goodness of fit test for time series”, unpublished manuscript, North Carolina State University.

Jennrich, R.I. (1969): “Asymptotic properties of nonlinear least squares estimators”, *Annals of Mathematical Statistics*, 40, 633-643.

Jurečkova, J. and Hallin, M. (1999): “Optimal tests for autoregressive models based on autoregression rank scores”, *The Annals of Statistics*, 27, 1385–1414.

Jurečkova, J. and Prochazka (1994): “Regression quantiles and trimmed least squares estimation in nonlinear regression models,” *Journal of Nonparametric Statistics*, 3, 201–222.

Khmaladze, E.V. (1981): “Martingale approach to the goodness of fit tests”, *Theory of Probability and Its Applications*, 26, 246-265.

Kim,T.H. and White,H. (2003): “Estimation, Inference, and Specification Testing for Possibly Misspecified Quantile Regressions”, in T. Fomby and R.C. Hill, eds., Maximum Likelihood Estimation of Misspecified Models: Twenty Years Later. New York: Elsevier, 107-132.

Koenker, R. and Bassett, G. (1978): “Regression quantiles”, *Econometrica*, 46, 33-50.

Koenker, R., and Machado, J.A. (1999): “Goodness of fit and related inference processes for quantile regression,” *Journal of the American Statistical Association*, 94, 1296-1310.

Koenker, R. and Park, B.J. (1996): “An interior point algorithm for nonlinear quantile regression”, *Journal of Econometrics*, 71, 265-285.

Koenker, R. and Xiao, Z. (2002): "Inference on the quantile regression process", *Econometrica*, 70, 1583–1612.

Koenker, R. and Xiao, Z. (2004): "Quantile autoregression", working paper.

Koul, H. L. (2002). *Weighted Empirical Processes in Dynamic Nonlinear Models.* 2nd ed, Lecture Notes in Statistics, Vol. 166, Springer.

Koul, H. L, and Ling, S. (2006): "Fitting an error distribution is some heteroscedastic time series models", forthcoming in *The Annals of Statistics.*

Koul, H.L., and Saleh, A.K.M.E (1994) "Autoregression quantile and related rank scores processes", *The Annals of Statistics*, 23, 670-689.

Koul, H.L., and Stute W. (1999): "Nonparametric model checks for time series", *The Annals of Statistics*, 27, 204-236.

Levental, S. (1989): "A uniform CLT for uniformly bounded families of martingale differences", *Journal of Theoretical Probability*, 2, 271-287.

Mukherjee, K. (1999): "Asymptotics of quantiles and rank scores in nonlinear time series", *Journal of Time Series Analysis*, 20(2), 173–192.

Nishiyama, Y. (2000): "Weak convergence of some classes of martingales with jumps", *The Annals of Probability*, 28, 685-712.

Pollard, D. (1991): "Asymptotics for least absolute deviation regression estimators", *Econometric Theory*, 7, 186-199.

Politis, D. N., Romano, J.P. and Wolf, M. (1999): *Subsampling.* Springer-Verlag, New York.

Portnoy, S. (1984): "Tightness of the sequence of empiric cdf processes defined from regression fractiles", in Robust and Nonlinear Time Series Analysis, eds. J. Franke, W. Hardle, and D. Martin, Springer-Verlag: New York.

Sakov, A. and Bickel, P. (2000): "An Edgeworth expansion for the m out of n bootstrap median", *Statistical and Probability Letters*, 49, 217-223.

Sharpe, W.F. (1988): "Determining a Fund's Effective Asset Mix", *Investment Management Review*, December, 59-69.

Sharpe, W.F. (1992): "Asset Allocation: Management Style and Performance Measurement", *Journal of Portfolio Management*, 18, 7-19.

Stinchcombe, M., and White, H. (1998): "Consistent specification testing with nuisance parameters present only under the alternative", *Econometric Theory*, 14, 295-325.

Stute, W. (1997): "Nonparametric model checks for regression", *The Annals of Statistics*, 25, 613-641.

van der Vaart, A.W., and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes.* Springer. New York.

Whang, Y-J. (2004): "Consistent specification tests for quantile regression models", forthcoming in Frontiers of Analysis and Applied Research: Essays in Honor of Peter C.B. Phillips, Cambridge University Press.

Whang, Y-J. (2005): "Smoothed Empirical Likelihood Methods for Quantile Regression Models", forthcoming in *Econometric Theory.*

Zheng, J.X. (1998): "A consistent nonparametric test of parametric regression models under conditional quantile restrictions", *Econometric Theory*, 14, 123-138.

Department of Economics

Indiana University

100 S. Woodlawn, Wylie Hall

Bloomington, IN 47405-7104, USA

E-mail: jescanci@indiana.edu

Departamento de Economía

Universidad Carlos III de Madrid

C./ Madrid, 126

28903 Getafe (Madrid), Spain

E-mail: Carlos.Velasco@uc3m.es

Table I:

Asymptotic Critical Values. Cramér-von Mises

	$m=9$			$m=15$			$m=21$		
ϵ	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.05	0.343	0.458	0.744	0.343	0.458	0.738	0.343	0.458	0.738
0.10	0.334	0.447	0.731	0.333	0.446	0.730	0.333	0.447	0.729
0.15	0.317	0.429	0.694	0.317	0.428	0.693	0.317	0.428	0.694
0.20	0.293	0.398	0.652	0.293	0.398	0.651	0.293	0.398	0.651
0.25	0.262	0.358	0.598	0.262	0.358	0.596	0.263	0.358	0.598
0.30	0.222	0.308	0.514	0.222	0.308	0.511	0.222	0.307	0.513
0.35	0.176	0.243	0.412	0.176	0.244	0.412	0.175	0.244	0.411
0.40	0.124	0.172	0.292	0.123	0.172	0.291	0.123	0.172	0.290
0.45	0.064	0.091	0.156	0.064	0.091	0.156	0.064	0.091	0.156
	$m=50$			$m=100$			$m=1000$		
ϵ	10%	5%	1%	10%	5%	1%	10%	15%	1%
0.05	0.343	0.458	0.737	0.342	0.458	0.737	0.342	0.458	0.736
0.10	0.334	0.445	0.726	0.333	0.446	0.727	0.333	0.445	0.726
0.15	0.317	0.429	0.692	0.317	0.428	0.693	0.317	0.429	0.692
0.20	0.292	0.397	0.651	0.293	0.397	0.652	0.292	0.397	0.653
0.25	0.263	0.358	0.597	0.263	0.358	0.597	0.262	0.358	0.597
0.30	0.222	0.307	0.512	0.222	0.307	0.511	0.222	0.307	0.511
0.35	0.176	0.244	0.411	0.175	0.244	0.412	0.175	0.244	0.411
0.40	0.123	0.172	0.291	0.123	0.172	0.290	0.123	0.172	0.291
0.45	0.064	0.091	0.156	0.064	0.091	0.156	0.064	0.091	0.156

Table II:

Asymptotic Critical Values. Kolmogorov-Smirnov

	$m = 9$			$m = 15$			$m = 21$		
ϵ	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.05	1.028	1.163	1.427	1.075	1.211	1.473	1.099	1.232	1.496
0.10	1.039	1.177	1.453	1.083	1.218	1.492	1.105	1.242	1.509
0.15	1.053	1.188	1.453	1.094	1.227	1.496	1.114	1.246	1.515
0.20	1.064	1.201	1.466	1.099	1.236	1.504	1.118	1.255	1.527
0.25	1.069	1.208	1.482	1.103	1.240	1.518	1.120	1.258	1.535
0.30	1.069	1.213	1.484	1.097	1.241	1.514	1.112	1.254	1.530
0.35	1.058	1.206	1.488	1.083	1.230	1.512	1.095	1.242	1.525
0.40	1.035	1.183	1.473	1.054	1.202	1.492	1.064	1.213	1.502
0.45	0.986	1.141	1.445	0.999	1.153	1.456	1.006	1.160	1.462
	$m = 50$			$m = 100$			$m = 1000$		
ϵ	10%	5%	1%	10%	5%	1%	10%	15%	1%
0.05	1.143	1.276	1.541	1.167	1.298	1.565	1.201	1.333	1.598
0.10	1.147	1.283	1.556	1.168	1.304	1.574	1.201	1.337	1.604
0.15	1.153	1.286	1.554	1.173	1.306	1.572	1.202	1.335	1.601
0.20	1.153	1.292	1.561	1.171	1.309	1.579	1.198	1.336	1.606
0.25	1.152	1.290	1.566	1.168	1.307	1.584	1.191	1.330	1.607
0.30	1.138	1.284	1.559	1.154	1.298	1.574	1.173	1.318	1.593
0.35	1.118	1.266	1.551	1.131	1.278	1.561	1.145	1.292	1.577
0.40	1.083	1.232	1.519	1.093	1.241	1.530	1.103	1.251	1.539
0.45	1.018	1.173	1.475	1.024	1.178	1.482	1.027	1.181	1.484

TABLE III: Empirical size and power. 5% of significance level. DGP1

DGP1		CvM_n			$KS_{n,1}$	$CvM_{n,1}$	$KS_{n,2}$	$CvM_{n,2}$
c_1	n	$k=9$	$k=10$	$k=11$				
0.0	**100**	4.8	6.3	7.5	5.2	5.5	5.0	5.7
	300	4.0	4.0	4.3	4.5	4.3	5.5	5.6
0.1	**100**	43.4	44.1	38.3	61.1	69.4	97.7	99.1
	300	98.2	97.3	97.0	98.3	98.7	100	100
0.2	**100**	81.0	78.4	69.8	93.3	95.4	100	100
	300	100	100	100	100	100	100	100
0.3	**100**	93.4	92.0	87.4	98.2	98.9	100	100
	300	100	100	100	100	100	100	100

TABLE IV: Empirical size and power. 5% of significance level. DGP2

DGP2		CvM_n			$KS_{n,1}$	$CvM_{n,1}$	$KS_{n,2}$	$CvM_{n,2}$
c_2	n	$k=3$	$k=4$	$k=5$				
0.0	**100**	5.1	5.1	6.1	5.2	5.9	5.2	5.5
	300	4.9	4.7	4.3	5.5	6.0	4.6	4.7
0.1	**100**	9.1	9.4	9.1	14.8	15.8	25.1	29.7
	300	22.1	21.6	20.8	36.5	41.4	73.9	78.4
0.2	**100**	23.8	23.7	23.6	40.0	44.5	73.6	80.1
	300	76.4	75.4	73.0	87.0	91.7	99.9	100
0.3	**100**	46.2	43.1	44.2	64.8	70.9	94.7	97.4
	300	97.2	97.5	96.7	98.7	99.1	100	100