Working Paper 97-41
Statistics and Econometrics Series 17
May 1997

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

CONSISTENT SPECIFICATION TESTING OF QUANTILE REGRESSION MODELS

Miguel A. Delgado and Manuel A. Domínguez*

Abstract

This paper introduces a specification testing procedure for quantile regression functions consistent in the direction of nonparametric alternatives. We consider test statistics based on a marked empirical process which does not require to estimate nonparametrically the true model. In general, the tests are not distribution free, but critical values can be consistently approximated using a residual based bootstrap. A small Monte Carlo experiment shows that the test works fairly well in practice.

Keywords:
Quantile regression, consistent specification test, marked empirical process, bootstrap.

*Delgado, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126 28903 Madrid. Spain. Ph: 34-1-624.98.04, Fax: 34-1-624.98.49, e-mail: delgado@est-econ.uc3m.es; Domínguez, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, 126 28903 Getafe, Madrid. Research funded by Spanish "Dirección General de Enseñanza Superior (DGES), reference number: PB95-0292.

CONSISTENT SPECIFICATION TESTING OF QUANTILE REGRESSION MODELS*

Miguel A. Delgado and Manuel A. Domínguez

Dept. of Statistics and Econometrics,

Universidad Carlos III de Madrid.

May 25, 1997

Abstract

This paper introduces a specification testing procedure for quantile regression functions consistent in the direction of nonparametric alternatives. We consider test statistics based on a marked empirical process which does not require to estimate nonparametrically the true model. In general, the tests are not distribution free, but critical values can be consistently approximated using a residual based bootstrap. A small Monte Carlo experiment shows that the test works fairly well in practice.

Keywords: Quantile regression; Consistent specification test; Marked empirical process; Bootstrap.

*Research funded by Spanish "Dirección General de Enseñanza Superior" (DGES), reference number: PB95-0292.

1. INTRODUCTION

This paper proposes a specification test of quantile regression models consistent in the direction of nonparametric alternatives.

Regression quantiles were introduced by Koenker and Bassett (1978) motivated by robustness considerations. Afterwards, quantile regression has also been proven very useful in applied research for describing conditional distributions, providing more accurate information on the relationship among the dependent and the conditioning random variables, than a mere conditional location estimate, like the conditional mean or median (see e.g. Koenker and Bassett (1982), Powell (1984, 1986), Granger et al. (1989), Chamberlain (1991) and Buchinsky (1994, 1995)).

In general, the same functional form is assumed for every quantile function. Such assumption seems very strong in practice when heteroskedasticity is present or when the underlying distribution is not standard. The proposed test statistics are based on a marked empirical process, where the marks depend on the quantile residuals of the model fitted consistently under the null hypothesis. The statistic is constructed in a similar way than those of Hong-zhy and Bing (1991), Su and Wei (1991), Delgado (1993), Diebold (1995), Andrews (1996) and Stute (1995). The true model needs not be estimated. Therefore, unlike other consistent specification testing procedures based on comparing the parametric estimator – consistent under the null hypothesis – with another a nonparametric estimate – consistent under both, the null and alternative hypotheses –, our test does not depend on the choice of a particular amount of smoothing. The test statistics is not, in general, distribution free. In order to implement the test, we also propose a residual based bootstrap procedure to approximate the critical values. A small simulation shows that the test works fairly well in practice.

The rest of the paper is organized as follows. Next Section introduces the testing procedure justifying its asymptotic properties. Section 3 presents the validity of bootstrap approximations. Section 4 reports the results of a small Monte Carlo experiment. Proofs are confined to Section 5.

2. TESTING PROCEDURE

Suppose we have observations $\{(Y_i, X_i), i = 1, ..., n\}$ independent and identically distributed as the $\mathbb{R} \times \mathbb{R}^d$-valued random vector $(\mathbf{Y}, \mathbf{X})$. Let $Q_\theta(\mathbf{Y} \mid \mathbf{X})$ be the θ conditional quantile function of $\mathbf{Y}$, i.e. $Q_\theta(\mathbf{Y} \mid \mathbf{X} = x) = \inf\left\{y : F_{\mathbf{Y}|\mathbf{X}}(y \mid \mathbf{X} = x) \geq \theta\right\}$, where $F_{\mathbf{Y}|\mathbf{X}}(. \mid .)$ is the conditional distribution of $\mathbf{Y}$ given $\mathbf{X}$. We are interested in testing the composite hypothesis

$$H_0 : \Pr\{Q_\theta(\mathbf{Y} \mid \mathbf{X}) = m_\theta(\mathbf{X}, \beta_0(\theta))\} = 1 \text{ some } \beta_0(\theta) \in B \subset \mathbb{R}^b,$$

and the alternative hypothesis, H_1, is the negation of H_0, where $m_\theta(\cdot,\cdot)$ is a known function, $\beta_0(\theta)$ is a vector of unknown parameters and $B \subset \mathbb{R}^b$ is the parameter space. Define $\Psi_\theta(z) = 1(z \leq 0) - \theta$, where $1(A)$ is the indicator function of the event A. Assuming that $\mathbf{Y}$ has a continuous distribution conditionally on $\mathbf{X}$, we can write

$$H_0 : \Pr\{E[\Psi_\theta(\mathbf{Y} - m_\theta(\mathbf{X}, \beta_0(\theta))) \mid \mathbf{X}] = 0\} = 1 \text{ some } \beta_0(\theta) \in B \subset \mathbb{R}^b.$$

Define $\varepsilon_\theta = \mathbf{Y} - m_\theta(\mathbf{X}, \beta_0(\theta))$. Then we can write the tautological quantile regression model,

$$Y_i = m_\theta(X_i, \beta_0(\theta)) + \varepsilon_{\theta i}, \quad i = 1, ..., n,$$

where $Q_\theta(\varepsilon_{\theta i} \mid X_i) = 0$.

Noting that

$$\Pr(g(\mathbf{X}) = 0) = 1 \Leftrightarrow E\left(g(\mathbf{X}) \prod_{k=1}^{d} 1(\mathbf{X}_k \leq x_k)\right) = 0 \text{ for all } x \in \mathbb{R}^d,$$

where $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_d)'$ and $x = (x_1, x_2, ..., x_d)'$, the null hypothesis can be equivalently expressed as

$$H_0 : E\left\{[\Psi_\theta(\mathbf{Y} - m_\theta(\mathbf{X}, \beta_0(\theta)))]\,\bar{\Delta}(x)\right\} = 0, \quad \text{all } x \in \mathbb{R}^d \text{ and some } \beta_0(\theta) \in B \subset \mathbb{R}^b, \tag{1}$$

where henceforth $\bar{\Delta}(x) = \prod_{k=1}^{d} 1(\mathbf{X}_k \leq x_k)$. We will base the test statistic on the sample analogue of the expectation in (1). Define

$$\bar{T}_{n\theta}^{0}(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Psi_{\theta i} \bar{\Delta}_i(x),$$

where $\Psi_{\theta i} = \Psi_\theta(\varepsilon_{\theta i})$ and $\bar{\Delta}_i(x) = \prod_{k=1}^{d} 1(X_{ik} \leq x_k)$, where $X_i = (X_{i1}, X_{i2}, ..., X_{id})$. Then, expression (1) suggests the following statistics for testing H_0,

$$CM_n \equiv \int_{-\infty}^{\infty} \left(\bar{T}_{n\theta}^{0}(x)\right)^2 dF_{n\mathbf{X}}(x) = n^{-1} \sum_{i=1}^{n} \bar{T}_{n}^{0}(X_i)^2. \quad (2)$$

$$KS_n \equiv \sup_{x \in \mathbb{R}^d} \left|\bar{T}_{n\theta}^{0}(x)\right| = \sup_{1 \leq i \leq n} \left|\bar{T}_{n\theta}^{0}(X_i)\right| \quad (3)$$

where $F_{n\mathbf{X}}(x)$ is the empirical distribution function of the regressors $\mathbf{X}$. CM_n and KS_n resemble in the spirit the Cramér-vonMises and Kolmogorov-Smirnov test statistics respectively. Although the KS_n is computationally more demanding, it could be more powerful than the CM_n against some alternatives (see Stute, 1995 for a discussion of the power of different functionals for the conditional mean specification test). This short of statistics has been used before for specification testing of regression models by Hong-zhy and Bin (1991), Su and Wei (1991), Delgado (1993), Diebold (1995) and Stute (1995) among others. Instead of $\bar{\Delta}_i(x)$, other continuous weight functions can also be considered as it has been proposed by Bierens (1982, 1990), De Jong and Bierens (1994), De Jong (1996) and Bierens and Ploberger (1997) for specification testing of continuous regression models. Notice that continuity of the underlying quantile regression model has not been imposed.

Althought CM_n and KS_n are based on a process on $D(-\infty, \infty)^d$, we can scale them into $D[0,1]^d$ by performing the quantile transformation

$$CM_n \equiv \int_{-\infty}^{\infty} T_{n\theta}^{0}(t)\ dF_{n\mathbf{U}}(t) = n^{-1} \sum_{i=1}^{n} T_{n}^{0}(U_i)^2,$$

$$KS_n \equiv \sup_{t \in [0,1]^d} \left|T_{n\theta}^{0}(t)\right| = \sup_{1 \leq i \leq n} \left|T_{n\theta}^{0}(U_i)\right|$$

where $\mathbf{U} = (\mathbf{U}_1, .., \mathbf{U}_d) = (F_{\mathbf{X}_1}(\mathbf{X}_1), .., F_{\mathbf{X}_d}(\mathbf{X}_d))$, $\{U_i : i = 1, .., n\}$ is the observed sample for $\mathbf{U}$, $F_{n\mathbf{U}}(t)$ is the empirical distribution function of $\mathbf{U}$ and

$$T_{n\theta}^{0}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Psi_{\theta i} \Delta_i(t),$$

where $\Delta_i(t) = \prod_{k=1}^{d} 1\left(F_{\mathbf{X}_k}(X_{ik}) \leq t_k\right)$, $t = (t_1, t_2, ..., t_d)'$, $t_k = F_{\mathbf{X}_k}(x_k)$ and $F_{\mathbf{X}_k}(x_k)$ is the unknown marginal distribution function of the regressor $\mathbf{X}_k$. Although the tests will always be computed using expressions (2) or (3), their asymptotic properties are easily discussed through $T_{n\theta}^0(t)$.

Under H_0, $E\left(1\left(\mathbf{Y} \leq m_\theta\left(\mathbf{X}, \beta_0(\theta)\right)\right) \mid \mathbf{X}\right) = \theta$ not depending on $\mathbf{X}$. Hence, the Bernoulli random variable $1\left(\mathbf{Y} \leq m_\theta\left(\mathbf{X}, \beta_0(\theta)\right)\right)$ is independent of $\mathbf{X}$. Let us define $\sigma_\theta = \sqrt{\theta(1-\theta)}$. Thus, when $d = 1$, by Donsker's invariance principle, $\sigma_\theta^{-1} T_n^0(t)$ converges in distribution to a standard Brownian motion on $D[0,1]$. This result is formally stated in the following proposition for the general case, $d \geq 1$.

Proposition 1 .- *Under* H_0,

$$T_{n\theta}^0(t) \to \sigma_\theta B^0(t) \ \textit{in distribution on the space } D[0,1]^d,$$

where $B^0(t) = \{B^0(t_1, t_2, ..., t_d),\ 0 < t_k < 1,\ k = 1, ..., d\}$ *is a Gaussian random process on* $[0,1]^d$ *centered at zero and with covariance structure given by*

$$\Sigma_0(t,s) = E\left(B^0(t) B^0(s)\right) = E\left(\prod_{k=1}^{d} 1\left(F_{\mathbf{X}_k}(\mathbf{X}_k) \leq \min(t_k, s_k)\right)\right)$$

where $s = (s_1, s_2, ..., s_d)'$.

Thus, under H_0, by the Continuous Mapping Theorem (CMT),

$$\begin{aligned} CM_n &\to CM_\infty \equiv \sigma_\theta^2 \int_{[0,1]^d} B^0(t)^2\, dt \text{ in distribution,} \\ KS_n &\to KS_\infty \equiv \sigma_\theta^2 \sup_{[0,1]^d} \left|B^0(t)\right| \text{ in distribution.} \end{aligned}$$

It is worth remarking that when independence among explanatory variables holds, $\Sigma_0(t,s) = \prod_{k=1}^{d} \min(t_k, s_k)$ and $B^0(t)$ is standard Brownian Motion in $D[0,1]^d$. Therefore, when $d = 1$, or $d > 1$ and the $\mathbf{X}'s$ are independent, the asymptotic distribution is known and tabulated. Hence, asymptotic tests can be implemented. Nevertheless, in the general case, the asymptotic distribution depends on the data generating process and the test will be more difficult to implement.

To test the composite hypothesis, the parameter $\beta_0(\theta)$ defined by

$$\beta_0(\theta) = \arg \min_{\beta \in B \subset \mathbb{R}^b} E\left\{\left(\left(m_\theta(\mathbf{X},\beta) - \mathbf{Y}\right)\Psi_\theta\left(\mathbf{Y} - m_\theta(\mathbf{X},\beta)\right)\right)\right\}. \tag{4}$$

can be estimated by its sample analogue

$$\hat{\beta}(\theta) = \arg \min_{\beta \in B} \frac{1}{n} \sum_{i=1}^{n} \left(m_\theta(X_i,\beta) - Y_i\right)\Psi_\theta\left(Y_i - m_\theta(X_i,\beta)\right). \tag{5}$$

Notice that the interpretation of $\beta_0(\theta)$ is different under the null and under the alternative hypothesis.

The asymptotic properties of $\hat{\beta}(\theta)$ have been studied under different conditions on the data generating process under the null. Koenker and Bassett (1978, 1982) and Bloomfield and Stieger (1983) consider the linear model. Phillips (1991) and Pollard (1991) introduced a different methodology. Amemiya (1982) discusses the properties of these sort of estimators for simultaneous equation models. Powell (1984, 1986) apply median and quantile regression to censored and truncated regression models. Oberhoffer (1982) proves consistency for the median in the nonlinear regression model. Recently, Weiss (1994) has obtained the asymptotic distribution of $\hat{\beta}(\theta)$ in general nonlinear dynamic models. It can be shown (see e.g. in Ruppert and Carroll, 1980 and Weiss, 1994 among others) that

$$\sum_{i=1}^{n} m_\theta^{(1)}\left(X_i,\hat{\beta}(\theta)\right)\Psi_\theta\left(Y_i - m_\theta\left(X_i,\hat{\beta}(\theta)\right)\right) = o_p(n^{1/2}), \tag{6}$$

holds for the solution to the problem (5), where $m_\theta^{(1)}(X_i,\beta) = \partial m_\theta(X_i,\beta)/\partial\beta$. Under different regularity conditions, it has been shown that $\hat{\beta}(\theta) = \beta_0(\theta) + O_p\left(n^{-1/2}\right)$ under H_0. For notational convenience, we concentrate on the linear in parameters case, $m_\theta(X_i,\beta_0(\theta)) = Z_i'\beta_0(\theta)$, where $Z_i = (1, X_i')'$. That is, we assume that an intercept is included in the model. Then, (6) becomes

$$\sum_{i=1}^{n} Z_i \hat{\Psi}_{\theta i} = o_p(n^{1/2}), \tag{7}$$

where $\hat{\Psi}_{\theta i} = \Psi_\theta\left(Y_i - Z_i'\hat{\beta}(\theta)\right)$.

When we replace the unknown $\beta_0(\theta)$ in $\bar{T}_n^0(x)$ by the estimator defined by (5), the statistic becomes

$$\bar{T}_{n\theta}^1(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{\Psi}_{\theta i} \bar{\Delta}_i(x).$$

Again, $\bar{T}_{n\theta}^1(x)$ and its quantile transformation $T_{n\theta}^1(t)$ are identical.

The following Proposition is useful in order to derive the asymptotic limiting process of $T_{n\theta}^1(t)$. We require the following conditions:

A.1 Let $F_{\varepsilon_\theta}(\varepsilon \mid \mathbf{X})$ be the conditional distribution of ε_θ given $\mathbf{X}$. Then, we assume

A.1.1 $F_{\varepsilon_\theta}(\varepsilon \mid \mathbf{X})$ has, at least, one bounded continuous derivative in a neighborhood of $\varepsilon = 0$, uniformly in x.

A.1.2 Define $f_{\varepsilon_\theta}(\varepsilon \mid \mathbf{X}) = dF_{\varepsilon_\theta}(\varepsilon \mid \mathbf{X})/d\varepsilon$. Then, $f_{\varepsilon_\theta}(0 \mid \mathbf{X}) = f_{\varepsilon_\theta}(0)$.

A.2 The regressors $\mathbf{X}$ have a continuous distribution such that $E\|\mathbf{X}\|^2 < \infty$.

Condition for **A.1.2** is satisfied when X_i and $\varepsilon_{\theta i}$ are independent as it is usually assumed in the literature related to quantile regression. This assumption is notationally convenient but it can easily relaxed as we shall discuss later. We shall also assume that conditions for $\sqrt{n}$-consistency of $\hat{\beta}(\theta)$ to $\beta_0(\theta)$ are satisfied.

Proposition 2 .- *Let $h(X,t)$ be a measurable function such that* $\sup_{t\in[0,1]^d} E\|h(\mathbf{X},t)\|^2 < \infty$. *Under H_0, **A.1-A.2**, uniformily in t,*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} h_i(t)\hat{\Psi}_{\theta i} - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h_i(t)\Psi_{\theta i} - f_{\varepsilon_\theta}(0)\ \frac{1}{n} \sum_{i=1}^{n} h_i(t) Z_i' \sqrt{n}\left(\hat{\beta}(\theta) - \beta_0(\theta)\right) = o_p(1), \tag{8}$$

where $h_i(t) = h(X_i,t)$.

Notice that if we choose $h_i(t) = Z_i$, the first summand in (8) vanishes asymptotically by condition (7), so we can write,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i \Psi_{\theta i} + f_{\varepsilon_\theta}(0)\ \frac{1}{n} \sum_{i=1}^{n} Z_i Z_i' \sqrt{n}\left(\hat{\beta}(\theta) - \beta_0(\theta)\right) = o_p(1).$$

Thus, from Proposition 2, and assuming $f_{\varepsilon_\theta}(0) > 0$, we can obtain the usual linearization of $\hat{\beta}(\theta)$, just replacing $h_i(t)$ by Z_i, i.e.,

$$\sqrt{n}\left(\hat{\beta}(\theta) - \beta_0(\theta)\right) = -\left[f_{\varepsilon_\theta}(0)\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i'\right]^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i \Psi_{\theta i} + o_p(1). \tag{9}$$

Therefore, substituting (9) in (8) and applying the Law of Large Numbers (LLN) when $h_i(t) = \Delta_i(t)$ we have, uniformly in t,

$$T_{n\theta}^1(t) = T_{n\theta}^0(t) - A(t)' R^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i \Psi_{\theta i} + o_p(1), \tag{10}$$

where $t \in [0,1]^d$, $R = E[Z_1' Z_1]$ and $A(t) = E[Z_1 \Delta_1(t)]$. From (10), it is easy to find the asymptotic covariance structure of $T_{n\theta}^1(t)$. The following theorem provides the limiting process of $T_{n\theta}^1(t)$ under H_0.

Theorem 1 .- *Under H_0,* ***A.1-A.2****,*

$$T_{n\theta}^1(t) \to \sigma_\theta B^1(t) \text{ in distribution on the space } D[0,1]^d,$$

where $B^1(t)$ is a Gaussian process centered at cero and covariance structure given by

$$\Sigma_1(t,s) = \Sigma_0(t,s) - A(t)' R^{-1} A(s).$$

When the quantile function is non linear in parameters, $\Sigma_1(t,s)$ will also depend on $\beta_0(\theta)$ and the model under the null hypothesis. In this case $A(t) = E\left[m^{(1)}(\mathbf{X}, \beta_0(\theta))\Delta_1(t)\right]$, and $R = E\left[m^{(1)}(\mathbf{X}, \beta_0(\theta))\, m^{(1)}(\mathbf{X}, \beta_0(\theta))'\right]$. If, in addition, we allow that $f_{\varepsilon_\theta}(0) \neq f_{\varepsilon_\theta}(0 \mid \mathbf{X})$, the covariance structure becomes

$$\begin{aligned}\Sigma_1(t,s) &= \Sigma_0(t,s) - \ddot{A}(t)' \ddot{R}^{-1} A(t) - A(t)' \ddot{R}^{-1} \ddot{A}(s) \\ &\quad + \ddot{A}(t)' \ddot{R}^{-1\prime} R\, \ddot{R}^{-1} \ddot{A}(s),\end{aligned}$$

where $\ddot{A}(t) = E\left[m^{(1)}(\mathbf{X}, \beta_0(\theta))\, f_{\varepsilon_\theta}(0 \mid \mathbf{X})\Delta_1(t)\right]$ and
$\ddot{R} = E\left[f_{\varepsilon_\theta}(0 \mid \mathbf{X}) m^{(1)}(\mathbf{X}, \beta_0(\theta))\, m^{(1)}(\mathbf{X}, \beta_0(\theta))'\right]$.

The test is based on the statistics,

$$\begin{aligned}CM_n^1 &\equiv \frac{1}{n}\sum_{i=1}^{n} \bar{T}_n^1(X_i)^2 = \frac{1}{n}\sum_{i=1}^{n} T_n^1(U_i)^2, \\ KS_n^1 &\equiv \sup_{1\le i\le n}\left|\bar{T}_n^1(X_i)\right| = \sup_{1\le i\le n}\left|T_n^1(U_i)\right|.\end{aligned}$$

Applying Theorem 1 and the continuous mapping theorem, the asymptotic distribution of CM_n^1 and KS_n^1 are immediately obtained as stated in the following Corollary.

Corollary 1 .- *Suppose **A.1-A.2** hold. Then under* H_0

$$CM_n^1 \xrightarrow{d} CM_\infty^1 \equiv \sigma_\theta^2 \int_{[0,1]^d} B^1(t)^2 \, dt$$
$$KS_n^1 \xrightarrow{d} KS_\infty^1 \equiv \sigma_\theta^2 \sup_{t \in [0,1]^d} \left| B^1(t) \right|.$$

Under H_1

$$\lim_{n\to\infty} \Pr\left\{CM_n^1 > c\right\} = \lim_{n\to\infty} \Pr\left\{KS_n^1 > c\right\} = 1 \text{ for all } c.$$

Remark 1 .- *When we test that the conditional quantile function is a constant, i.e.* $H_0 : \Pr\{Q(\mathbf{Y} \mid \mathbf{X}) = \beta_0(\theta)\} = 1$, *for some unknown scalar* $\beta_0(\theta)$ *and* $d = 1$ *or* $d > 1$ *but the regressors are independent, then* $R = 1$ *and* $A(t) = t$ *and we have that* $B^1(t)$ *in Theorem 1 is the standard Brownian Bridge on* $D[0,1]^d$*, and* CM_n^1 *and* KS_n^1 *share the same asymptotic null distribution as the Cràmer-v.Mises and Kolmogorov-Smirnov statistics employed for testing goodness of fit of the parametric distribution function.*

In general, both CM_∞^1 and KS_∞^1 are not distribution free. However, critical values can be consistently approximated by bootstrap, as discussed in the following Section.

3. BOOTSTRAP TEST

As it has been shown, the asymptotic null distribution of any statistic based on $\bar{T}_{n\theta}^0(x)$ or $\bar{T}_{n\theta}^1(x)$ depends, in general, on certain characteristics of the data generating process. In order to implement the test in practice, we propose a residual based näive bootstrap procedure assuming independence between errors and regressors.

First, we discuss the bootstrap approximation for the simple hypothesis in which $\beta(\theta)$ is known and errors, $\varepsilon_{\theta i}$, are observable. Let $\mathcal{X} = \{(X_i, \varepsilon_{\theta i}) : i = 1, .., n\}$ be the observed sample of the regressors and the error term. Suppose $\{\varepsilon_{\theta i}^* : i = 1, .., n\}$ is a random sample drawn from a multinomial distribution that puts equal weight on the observed errors $\{\varepsilon_{\theta i} : i = 1, .., n\}$. Let us define $\theta_n = n^{-1} \sum_{i=1}^n 1(\varepsilon_{\theta i} \leq 0)$.

Henceforth, we will use standard bootstrap notation, i.e., we define $E^*(\cdot) = E(\cdot \mid \mathcal{X})$, $Var^*(\cdot) = Var(\cdot \mid \mathcal{X})$ and $\Pr^*(\cdot) = \Pr(\cdot \mid \mathcal{X})$. Define $\Psi^*_{\theta i} = 1(\varepsilon^*_{\theta i} < 0) - \theta_n$. Next Proposition guarantees that the distribution of $T^0_n(t)$ can be approximated by the one of its bootstrap analogue

$$T^{0*}_{n\theta}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Psi^*_{\theta i} \Delta_i(t).$$

Proposition 3 *With probability one*

$$T^{0*}_{n\theta}(t) \to \sigma_\theta B^0(t) \text{ in distribution in the space } D[0,1]^d.$$

When $\beta(\theta)$ is unknown, errors are unobserved and we will resample from the residuals. Write $\hat{\varepsilon}_{\theta i} = Y_i - Z_i'\hat{\beta}(\theta)$, $i = 1, .., n$. Now we define $\hat{\theta}_n = n^{-1} \sum_{i=1}^{n} 1(\hat{\varepsilon}_{\theta i} \leq 0)$. By (7), taking into account that a constant is present, $\hat{\theta}_n = \theta + o_p(n^{-1/2})$. Now let $\hat{\varepsilon}^*_{\theta 1}, .., \hat{\varepsilon}^*_{\theta n}$ be random samples drawn from a multinomial distribution that puts equal weight on the residuals $\{\hat{\varepsilon}_{\theta i} : i = 1, .., n\}$. Define $\hat{\Psi}^*_{\theta i} = 1(\hat{\varepsilon}^*_{\theta i} < 0) - \hat{\theta}_n$. By construction we have, $E^*\left[\hat{\Psi}^*_{\theta i}\right] = 0$ both under H_0 and H_1, where now $\mathcal{X} = \{(Y_i, X_i), i = 1, ..., n\}$ is the observed sample. Define $Y_i^* = Z_i'\hat{\beta}(\theta) + \hat{\varepsilon}^*_{\theta i}$ and

$$\hat{\beta}^*(\theta) = \arg\min_{\beta \in B} \frac{1}{n} \sum_{i=1}^{n} (Z_i'\beta - Y_i^*)(1(Y_i^* - Z_i'\beta \leq 0) - \theta)$$

De Angelis et al (1993) and Hahn (1995) have shown that the bootstrap distribution of $\sqrt{n}\left(\hat{\beta}^*(\theta) - \hat{\beta}(\theta)\right)$ converges, in probability, to the distribution of $\sqrt{n}\left(\hat{\beta}(\theta) - \beta_0(\theta)\right)$ for the linear model and fixed regressors. However, it is straightforward to generalize the proof of Theorem 1 in Hahn (1995) in order to allow for stochastic regressors.

The bootstrap analogue of $T^1_{n\theta}(t)$ is given by

$$\hat{T}^{1*}_{n\theta}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{\Psi}^*_{\theta i} \Delta_i(t),$$

where $\tilde{\Psi}^*_{\theta i} = 1(\tilde{\varepsilon}^*_{\theta i} < 0) - \hat{\theta}_n$ and $\tilde{\varepsilon}^*_{\theta i} = Y_i^* - Z_i'\hat{\beta}^*(\theta)$, $i = 1, .., n$, the residuals of the bootstrap estimation. Define $\hat{T}^{0*}_{n\theta}(t) = n^{-1/2} \sum_{i=1}^{n} \hat{\Psi}^*_{\theta i} \Delta_i(t)$.

Proposition 4 .- *Let $h(X,t)$ be as in Proposition 2. Then, under H_0,* ***A.1-A.2****, uniformly in t,*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_i(t)\,\tilde{\Psi}^*_{\theta i} - \frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_i(t)\,\hat{\Psi}^*_{\theta i} - f_{\varepsilon_\theta}(0)\left(\frac{1}{n}\sum_{i=1}^{n} h_i(t)\,Z_i'\right)\sqrt{n}\left(\hat{\beta}^*(\theta)-\hat{\beta}(\theta)\right)$$
$$+d_n\left(t,\hat{\beta}^*(\theta)-\hat{\beta}(\theta)\right) = o_{p^*}(1),$$

in probability, where

$$d_n\left(t,\hat{\beta}^*(\theta)-\hat{\beta}(\theta)\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{F_n\left(Z_i'\left(\hat{\beta}^*(\theta)-\hat{\beta}(\theta)\right)\right) - F\left(Z_i'\left(\hat{\beta}^*(\theta)-\hat{\beta}(\theta)\right)\right)\right.$$
$$\left. -F_n(0)+F(0)\right\}\ h_i(t)$$

and

$$V_n^* = o_{p^*}(1) \text{ in probability} \Leftrightarrow \Pr\left(\|V_n^*\| > \delta \mid \mathcal{X}\right) \xrightarrow{p} 0 \text{ for all } \delta > 0.$$

This linearization is the bootstrap analogue of Proposition 2. Notice that now a bias term $d_n\left(t,\hat{\beta}^*(\theta)-\hat{\beta}(\theta)\right)$ appears. Next Proposition shows that this bias can be asymptotically approximated by a term which is constant conditionally on $\mathcal{X}$ and unconditionally vanishing.

Proposition 5 *Let $0<\gamma<1/2$. Then*

$$\left|d_n\left(t,\hat{\beta}^*(\theta)-\hat{\beta}(\theta)\right)\right| \le \sup_{\|\xi_n\|\le n^{-1/2+\gamma}} \left|d_n\left(t,\xi_n\right)\right| + o_p^*(1),$$

where

$$\sup_{\|\xi_n\|\le n^{-1/2+\gamma}} \left|d_n\left(t,\xi_n\right)\right| = o_p(1).$$

Therefore, just mimicing steps in expression (10), we write its bootstrap analogue,

$$\hat{T}^{1*}_{n\theta}(t) - \hat{T}^{0*}_{n\theta}(t) + \frac{1}{n}\sum_{i=1}^{n} Z_i'\Delta_i(t)\left[\frac{1}{n}\sum_{i=1}^{n} Z_iZ_i'\right]^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i\Psi^*_{n\theta} + \check{d}_n(t) = o_{p^*}(1) \quad (11)$$

where $\check{d}_n(t) = o_p(1)$, which is used to prove the following theorem.

Theorem 2 .- *Under H_0,* ***A.1-A.2****, in probability,*

$$\hat{T}^{1*}_{n\theta}(t) \to \sigma_\theta\, B^1(t) \text{ in distribution, in the space } D\,[0,1]^d,$$

Alternatively a smooth bootstrap procedure can be used, where the samples are drawn from a smooth nonparametric estimator of the density of the error term based on $\{\hat{\varepsilon}_i,\ i = 1,..,n\}$. It has been shown in De Angelis et al (1993) that smooth bootstrap estimates of quantiles, work better than those based on the näive bootstrap. It is straightforward to check that Theorem 2 also holds for the smooth bootstrap. Higher order comparison of the different test are outside the scope of this paper. However, in Monte-Carlo experiments we have performed, the test based on smooth bootstrap seems to be quite sensitive to the choice of the amount of smoothing and it does not perform better than the näive bootstrap.

4. SIMULATIONS

In the simulations, we consider the following design,

$$Y_i = \beta_{01} + \beta_{02}X_i + \beta_{03}X_i^2 + \beta_{04}X_i^3 + \varepsilon_i \quad i = 1,..,n, \tag{12}$$

for different values of β_{0j}, $j = 1,2,3,4$ and $n = 30, 50, 100$. The regressors X_i are independently distributed $U(0,1)$ and the errors are generated such that $Q_{0.5}(\varepsilon_i \mid X_i) = 0$. Two error distributions have been proved, $\varepsilon_i = u_i$ and $\varepsilon_i = (\exp(u_i) - 1) \,/\, (\exp(2) - \exp(1))$ (a standarized lognormal), where u_i are distributed iid $N(0,1)$.

Tables I to III show the proportion of rejections on 1000 replications for each model, sample size and error distribution when the CM_n test is applied.

In Table I we report the proportion of rejections using the asymptotic critical values for testing significance of the explanatory variable, that is, $H_0 : Q_{0.5}(\mathbf{Y} \mid \mathbf{X}) = \beta_{01}$. The proportion of rejections under H_0 is quite close to the theoretical size of the test, even for the smallest sample sizes. Note that, under H_0, the statistic is identical for both error distributions. We study the power of the test when the true model is a linear model, $\beta_{02} \neq 0$, but $\beta_{03} = \beta_{04} = 0$. The proportion of rejections grows fast, both with sample size and with the value taken by β_{02}.

Tables II and III illustrate the behavior of the bootstrap test. Here, the coefficients $\beta(\theta)$ have been estimated using the algorithm proposed by Koenker and D'Orey

(1987). In Table II, the null hypothesis consists of the linear model $Q_{0.5}(\mathbf{Y} \mid \mathbf{X}) = \beta_{01}+\beta_{02}\mathbf{X}$. The empirical size of the bootstrap test approximate very well the nominal size and it is not affected by the value taken by the parameters β_{01}, β_{02}. We also study the empirical power when the true model is quadratic ($\beta_{0i} \neq 0$, $i = 1, 2, 3$, but $\beta_{04} = 0$). The power of the test also grows fast with the sample size and with the value taken by β_{03}. Since $\mathbf{X}^2$ is much smaller than $\mathbf{X}$, the alternative hypothesis will be quite close to the null, except for the biggest value of β_{03}. Such conclusions also hold for Table III, where the null hypothesis is quadratic $Q_{0.5}(\mathbf{Y} \mid \mathbf{X}) = \beta_{01} + \beta_{02}\mathbf{X}+\beta_{03}\mathbf{X}^2+\beta_{04}\mathbf{X}^3$ while the data is generated by (12) with $\beta_{0i} \neq 0$, $i = 1, 2, 3$ and β_{04} is equal or different to 0.

5. MATHEMATICAL APPENDIX.

In this Section we prove the Propositions and Theorems stated in previous Sections, which are based on some technical Lemmas proved in the Lemmatas.

Proof of Proposition 1

Since $\{\Psi_{\theta i}\Delta_i(t) : i = 1, .., n\}$ are *iid* with zero expectation and for $t, s \in [0,1]^d$, $t \neq s$, $\Delta_i(t)\Delta_i(s) = \Delta_i(\min(t, s))$, then $E\left(\Psi_{\theta i}^2\Delta_i(t)\Delta_i(s)\right) = \sigma_\theta^2\Sigma_0\left(t, s\right)$ and applying the Levy Central Limit Theorem, the finite dimensional distribution of $T_{n\theta}^0(t)$ converges to a normal with zero mean and covariance given by Σ_0.

Let $D_1 = (s, t] = \times_{j=1}^d (s_j, t_j]$, $D_2 = (s', t'] = \times_{j=1}^d (s'_j, t'_j]$ be two neighbors intervals in $[0,1]^d$, i.e., they abut and for some $j \in \{1, 2, .., d\}$, and they have the same jth-face, $\times_{k\neq j}(s_k, t_k] = \times_{k\neq j}(s'_k, t'_k]$. Let $W_n(t)$ be any empirical random process on $D[0,1]^d$. Let define

$$W_n(D_1) = \sum_{e_1=0}^{1}\cdots\sum_{e_d=0}^{1}(-1)^{d-\sum_p e_p}\, W_n\left(s_1 + e_1(t_1 - s_1), \cdots, s_d + e_d(t_d - s_d)\right). \quad (13)$$

Expression (2.1.8) in Gaenssler and Stute (1979) assures that a sufficient condition for tightness in $D[0,1]^d$ is

$$\Pr\left(\left|W_n\left(D_1\right)\right| > \delta; \left|W_n\left(D_2\right)\right| > \delta\right) \leq K\delta^{-a}\left(\mu\left(D_1 \cup D_2\right)\right)^b \quad (14)$$

where $\mu(\cdot)$ is an arbitrary finite measure with continuous marginals, $b > 1$, $K \geq 0$, and a are arbitrary constants. Using Markov inequality, a sufficient condition for (14) is

$$E\left(|W_n(D_1)|\,|W_n(D_2)|\right)^2 \leq K\left(\mu\left(D_1 \cup D_2\right)\right)^b. \tag{15}$$

Without loss of generality we will prove tightness for $d = 2$. For our process, expression (13) can be written as

$$\begin{aligned} T^0_{n\theta}(D_1) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Psi_{\theta i}\left(\Delta_i(t_1,t_2) - \Delta_i(s_1,t_2) - \Delta_i(t_1,s_2) + \Delta_i(s_1,s_2)\right) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Psi_{\theta i}\,\Delta_i(D_1), \end{aligned}$$

where $\Delta_i(D_j) = 1(X_i \in D_j)$. In the same way we write

$$T^0_{n\theta}(D_2) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Psi_{\theta i}\,\Delta_i(D_2).$$

Lemma 5.1 in Stute (1995) assures that if $\{(\alpha_i,\beta_i)\}_{i=1}^n$ are n *iid* square integrable random vectors with $E(\alpha_1) = E(\beta_1) = 0$, then

$$E\left(\left(\sum_{i=1}^{n}\alpha_i\right)^2\left(\sum_{i=1}^{n}\beta_i\right)^2\right) \leq nE(\alpha_1^2\beta_1^2) + 3n(n-1)E(\alpha_1^2)E(\beta_1^2).$$

Now, if we call $\alpha_i = \Psi_{\theta i}\,\Delta_i(D_1)$ and $\beta_i = \Psi_{\theta i}\,\Delta_i(D_2)$ and taking into account that for our expression $\alpha_1^2\beta_1^2 = 0$ we have

$$\begin{aligned} E\left(\left(T^0_{n\theta}(D_1)\right)^2\left(T^0_{n\theta}(D_2)\right)^2\right) &\leq \frac{3n(n-1)}{n^2}E\left(\Psi_{\theta 1}\,\Delta_1(D_1)\right)^2 E\left(\Psi_{\theta 1}\,\Delta_1(D_2)\right)^2 \\ &\leq K\,\left(\Pr\left(X_1 \in D_1 \cup D_2\right)\right)^2. \end{aligned}$$

Thus, (15) holds and the proof is completed. ■

Proof of Proposition 2

For the sake of brevity, we define $\beta_\theta = \beta_0(\theta)$, $\hat{\beta}_\theta = \hat{\beta}(\theta)$ and

$$G(\delta) = \left\{\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}h_i(t)\hat{\Psi}_{\theta i} - \frac{1}{\sqrt{n}}\sum_{i=1}^{n}h_i(t)\Psi_{\theta i} - f_{\varepsilon_\theta}(0)\left(\frac{1}{n}\sum_{i=1}^{n}h_i(t)Z_i'\right)\sqrt{n}\left(\hat{\beta}_\theta - \beta_\theta\right)\right| > \delta\right\}.$$

Applying the definition of the function $\Psi_{\theta i}$,

$$G(\delta) = \left\{\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}h_i(t)\left[1\left(\varepsilon_{\theta i} \leq Z_i'\left(\hat{\beta}_\theta - \beta_\theta\right)\right) - 1\left(\varepsilon_{\theta i} \leq 0\right) - f_{\varepsilon_\theta}(0)Z_i'\left(\hat{\beta}_\theta - \beta_\theta\right)\right]\right| > \delta\right\}.$$

Let $0 < \gamma < 1/10$. Then

$$\Pr\{G(\delta)\} \le \Pr\left\{\left|\hat{\beta}_\theta - \beta_\theta\right| > n^{\gamma-1/2}\right\} + \Pr\left\{G(\delta) \text{ and } \left|\hat{\beta}_\theta - \beta_\theta\right| \le n^{\gamma-1/2}\right\}.$$

The first summand tends to 0 by $\hat{\beta}_\theta$ $\sqrt{n}$-consistency, and the second converges to 0 by Lemma 2. ■

Proof of Theorem 1

As in the proof of Proposition 1, we must show convergence of finite dimensional distributions and tightness. For the convergence of finite dimensional distributions, apply the Central Limit Theorem to the right hand side of expression (10), as it was done in the proof of Proposition 1. In order to prove tightness, note that expression (10) is the addition of two processes. By Proposition 1 the first summand is tight. Tightness of the second term follows immediately because the indexing deterministic function is continuous, nondecreasing and bounded. Thus, (10) is the addition of two tight processes. A process is said to be tight if there exists a compact set of the sample space, where the process is evaluated in, with arbitrary high probability, uniformly in n. Let K_1, K_2 the compact sets where the first and the second summand of (10) are evaluated in, with arbitrary high probability. By Tychonoff Theorem (See Dudley, 1989, th 2.2.8), the set $K_3 = \{k = (k_1, k_2) : k_1 \in K_1, k_2 \in K_2\}$ is compact with the product topology. Now because the addition is a continuous operator and any continuous transformation preserves compactness, the set $K_4 = \{k = k_1 + k_2 : (k_1, k_2) \in K_3\}$ is a compact set. Thus, the process in (10) is evaluated in the compact set K_4 with arbitrary large probability uniformily in n, and tightness follows. ■

Proof of Proposition 3

As in the proof of Proposition 1, we show weak convergence in $D[0,1]^d$ by verifying convergence of the finite dimensional distribution and tightness. First, we will prove the convergence of the finite dimensional distributions. Note that $T_n^{0*}(t)$ can be expressed as

$$\begin{aligned} T_{n\theta}^{0*}(t) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Psi_{\theta i}^{*} \Delta_i(t) \\ &= a(t)\, b_n(t)\, c_n^{*}(t)\, 1\left(F_{n\mathbf{U}}(t) > 0\right) \end{aligned}$$

where

$$a(t)=\sigma_\theta\sqrt{F_{\mathbf{U}}(t)},\ b_n(t)=\frac{\hat{\sigma}_{\theta_n}}{\sigma_\theta}\sqrt{\frac{F_{n\mathbf{U}}(t)}{F_{\mathbf{U}}(t)}},\ c_n^*(t)=\frac{1}{\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}}\sum_{i=1}^n\Psi_{\theta i}^*\Delta_i(t),$$

and $\hat{\sigma}_{\theta_n}=Var^*\left(\Psi_{\theta i}^*\right)=n^{-1}\sum_{i=1}^n\left(1\left(\varepsilon_i\le 0\right)-\theta_n\right)^2=\theta_n\left(1-\theta_n\right)$. Now suppose $t>0$. Then $F_{n\mathbf{U}}(t)\overset{a.s.}{\to}F_{\mathbf{U}}(t)>0$ by the Glivenko-Cantelli theorem and $1\left(F_{n\mathbf{U}}(t)>0\right)\overset{a.s.}{\to}1$. Thus, applying Lemma 6, it suffices to consider the convergence of $a(t)\ b_n(t)\ c_n^*(t)$, where $a(t)$ is constant for fixed t, $b_n(t)$ is also deterministic conditionally on the sample, and tends $a.s.$ to 1 uniformly in t by the Glivenko-Cantelli theorem and the strong consistency of $\hat{\sigma}_{\theta_n}$. Let define

$$\xi_{in}^*=\frac{\Delta_i(t)\Psi_{\theta i}^*}{\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}}.$$

Thus, $c_n^*(t)=\sum_{i=1}^n\xi_{in}^*$. It is easy to check that $E^*\left(\xi_{in}^*\right)=0\ a.s.$, $\sum_{i=1}^n Var^*\left(\xi_{in}^*\right)=1\ a.s.$ and ξ_{in}^*, ξ_{jn}^* are independent for $i\neq j\ a.s.$. Thus, $\{\xi_{in}^*, i=1,..,n;n=1,2,...\}$ satisfies the conditions of triangular arrays almost surely. Now we check the Lindeberg condition.

$$\begin{aligned}
\sum_{i=1}^n E^*\left(\xi_{in}^{*2}\,1\left(\left|\xi_{in}^*\right|>\delta\right)\right) &= \frac{1}{\hat{\sigma}_{\theta_n}^2 nF_{n\mathbf{U}}(t)}\sum_{i=1}^n E^*\left(\Psi_{\theta i}^{*2}\,1\left(\left|\frac{\Delta_i(t)\Psi_{\theta i}^*}{\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}}\right|>\delta\right)\right)\Delta_i(t)\\
&\le \frac{1}{\hat{\sigma}_{\theta_n}^2 F_{n\mathbf{U}}(t)}\sup_i E^*\left(\Psi_{\theta i}^{*2}\,1\left(\left|\Delta_i(t)\Psi_{\theta i}^*\right|>\delta\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}\right)\right)\\
&\le \frac{1}{\hat{\sigma}_{\theta_n}^2 F_{n\mathbf{U}}(t)}\sup_i E^*\left(1\left(\left|\Delta_i(t)\Psi_{\theta i}^*\right|>\delta\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}\right)\right)\\
&\le \frac{1}{\hat{\sigma}_{\theta_n}^2 F_{n\mathbf{U}}(t)}E^*\left(1\left(\left|\Psi_{\theta 1}^*\right|>\delta\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}\right)\right).
\end{aligned}$$

Now

$$\begin{aligned}
1\left(\left|\Psi_{\theta i}^*\right|>\delta\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}\right) &= 1\left(\varepsilon_{in}^*<0\right)1\left((1-\theta)>\delta\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}\right)\\
&\quad +1\left(\varepsilon_{in}^*>0\right)1\left(\theta>\delta\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}\right)
\end{aligned}$$

so

$$\begin{aligned}
\sum_{i=1}^n E^*\left(\xi_{in}^{*2}\,1\left(\left|\xi_{in}^*\right|>\delta\right)\right) &\le 1\left((1-\theta)>\delta\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}\right)+1\left(\theta>\delta\hat{\sigma}_{\theta_n}\sqrt{nF_{n\mathbf{U}}(t)}\right)\\
&\to 0\ \ a.s,
\end{aligned}$$

and we can conclude that the triangular array satisfies the Lindeberg's condition for almost every sample $\mathcal{X}$. Thus, by the CLT for triangular arrays we have, for fixed t

$$c_n^*(t) \xrightarrow{d} N(0,1), \quad a.s.$$

So by Lemma 6, for t fixed,

$$\hat{\sigma}_{\theta_n}^{-1} T_{n\theta}^{0*}(t) \xrightarrow{d} N(0, F_{\mathrm{U}}(t)) \quad a.s.$$

Noting that $E^*\left(T_{n\theta}^{0*}(t)T_{n\theta}^{0*}(s)\right) = \hat{\sigma}_{\theta_n} F_{n\mathrm{U}}\left(\min(t,s)\right)$, almost sure convergence of finite dimensional distributions follows.

For proving tightness, we use the same procedure followed by Stute, Manteiga and Presedo (1996) in the proof of their Lemma 4.3. That is, we will mimic conditionally on the sample all the steps given in the proof of Proposition 1. Using the same notation as in the referred Proposition,

$$\begin{aligned}
\hat{\sigma}_\theta^{-1} T_{n\theta}^{0*}(D_1) &= \hat{\sigma}_\theta^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \Psi_{\theta i}^* \left(\Delta_i(t_1,t_2) - \Delta_i(s_1,t_2) - \Delta_i(t_1,s_2) + \Delta_i(s_1,s_2)\right) \\
&= \hat{\sigma}_\theta^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \Psi_{\theta i}^* \Delta_i(D_1),
\end{aligned}$$

and in the same way we express

$$T_{n\theta}^{0*}(D_2) = \hat{\sigma}_\theta^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \Psi_{\theta i}^* \Delta_i(D_2),$$

where $\Delta_i(D_j)$ is not random conditionally on the sample. Now

$$\begin{aligned}
& E^*\left\{\hat{\sigma}_\theta^{-4} \left(T_{n\theta}^{0*}(D_1)\right)^2 \left(T_{n\theta}^{0}(D_2)\right)^2\right\} \\
= \;& \hat{\sigma}_\theta^{-4} E^* \left\{ \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \Psi_{\theta i}^* \Delta_i(D_1)\right)^2 \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \Psi_{\theta i}^* \Delta_i(D_2)\right)^2 \right\} \\
= \;& \frac{1}{\hat{\sigma}_\theta^4 n^2} \sum_{i=1}^n \sum_{j=1}^n E^*\left(\Psi_{\theta i}^{*2} \Psi_{\theta j}^{*2}\right) \left(\Delta_i(D_1)\Delta_j(D_2) + \Delta_i(D_1)\Delta_i(D_2)\Delta_j(D_1)\Delta_j(D_2)\right) \\
& + \Phi_n, \qquad (16)
\end{aligned}$$

where Φ_n includes terms containing, at least, one not repeated subindex. Thus $\Phi_n = 0$ because $E^*\left(\Psi_{\theta i}^* \Psi_{\theta j}^* \Psi_{\theta k}^* \Psi_{\theta l}^*\right) = 0$ for $i \neq j \neq k \neq l$. Now, taking into account that, by construction, $\Delta_i(D_1)\Delta_i(D_2) = 0$ and the independence among $\Psi^*_{\theta i}$ and $\Psi^*_{\theta j}$ for any $i \neq j$, (16) can be written as

$$\frac{1}{\hat{\sigma}_\theta^4 n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E^*\left(\Psi_{\theta i}^{*2}\right) E^*\left(\Psi_{\theta j}^{*2}\right)\Delta_i(D_1)\Delta_j(D_2) \leq \left(\frac{1}{n}\sum_{i=1}^{n} 1(X_i \in D_1 \cup D_2)\right)^2 \stackrel{a.s.}{\to} \Pr\left(\mathbf{X} \in D_1 \cup D_2\right)^2, \quad (17)$$

which is a sufficient condition for the tightness condition (15), conditionally on the sample, and the proof is completed. ■

Proof of Proposition 4.

Define $\hat{\beta}^*_\theta = \hat{\beta}^*(\theta)$ and $G^*(\delta)$ the bootstrap analogue of $G(\delta)$, i.e.,

$$G^*(\delta) = \left\{\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_i(t)\hat{\Psi}^*_{\theta i} - \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_i(t)\Psi^*_{\theta i} + f_{\varepsilon_\theta}(0)\,\frac{1}{n}\sum_{i=1}^{n} h_i(t)Z_i'\,\sqrt{n}\left(\hat{\beta}^*_\theta - \hat{\beta}_\theta\right)\right) + d_n(t, \hat{\beta}^*_\theta - \hat{\beta}_\theta)\right| > \delta\right\}.$$

Then

$$\Pr{}^*\left\{G^*(\delta)\right\} \leq \Pr{}^*\left\{\left|\hat{\beta}^*_\theta - \hat{\beta}_\theta\right| > n^{\gamma-1/2}\right\} + \Pr{}^*\left\{G^*(\delta) \text{ and } \left|\hat{\beta}^*_\theta - \hat{\beta}_\theta\right| \leq n^{\gamma-1/2}\right\}$$

but for any $\gamma > 0$, the first summand tends to 0 in probability by bootstrap $\sqrt{n}$-consistency proved by De Angelis et al (1993) and Hahn (1995). For any $\gamma < 1/10$, the second converges to 0 by Lemma 4 and the proof is completed. ■

Proof of Proposition 5

$$\begin{aligned}\left|d_n(t, \hat{\beta}^*_\theta - \hat{\beta}_\theta)\right| &\leq 1\left(\left|\hat{\beta}^*_\theta - \hat{\beta}_\theta\right| \leq n^{\gamma-1/2}\right) \sup_{\|\xi_n\| \leq n^{\gamma-1/2}} |d_n(t, \xi_n)| \\ &\quad + 1\left(\left|\hat{\beta}^*_\theta - \hat{\beta}_\theta\right| > n^{\gamma-1/2}\right)\left|d_n(t, \hat{\beta}^*_\theta - \hat{\beta}_\theta)\right| \\ &= \sup_{\|\xi_n\| \leq n^{\gamma-1/2}} |d_n(t, \xi_n)| + o_{p^*}(1),\end{aligned}$$

by the $\sqrt{n}$-consistency of the estimator.

Now

$$\sup_{\|\xi_n\| \leq n^{\gamma-1/2}} |d_n(t, \xi_n)| = o_p(1),$$

by Lemma 1.

Proof of theorem 2.

We start with expression (11) which, applying Proposition 4, is a valid expansion of $\hat{T}_n^{1*}(t)$ in probability, conditionally on the sample. The term $\check{d}_n(t)$ is negligible by Lemma 5. Now convergence of finite dimensional distributions of this expression is straightforward applying Lindeberg-Feller CLT as in the proof of Proposition 3 and Theorem 1. By Lemma 3, tightness is immediate if we show that the process is the sum of two tight processes. The proof of tightness for the first is identical to the proof of tightness of Proposition 3. Tightness of the second is immediate, and the proof is completed. ■

LEMMATA A

These Lemmata are applied in the proofs of results in Section 2. Henceforth, for the sake of presentation, we write $\varepsilon_i = \varepsilon_{\theta i}$, $F(\cdot) = F_{\varepsilon_\theta}(\cdot)$ and $f(\cdot) = f_{\varepsilon_\theta}(\cdot)$.

Lemma 1

Let $\check{\beta}_n - \beta$ be a sequence such that

$$\check{\beta}_n - \beta = \xi_n = O(n^{\gamma-1/2})$$

for any $\gamma \in (0, 1/10)$. If **A.1, A.2** are satisfied then

$$\sup_{|\xi_n| \le n^{\gamma-1/2}} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h_i(t) \left\{ 1\left(\varepsilon_i \le Z_i'\xi_n\right) - F\left(Z_i'\xi_n\right) - 1\left(\varepsilon_i \le 0\right) + F(0) \right\} \right| = o_p(1) \quad (18)$$

Proof. For simplicity, we prove it for X random scalar. We avoid to take the supremum on a infinite set, using the same procedure in Boldin (1982). He considers a greed of points in the range of variation $(-n^{\gamma-1/2}, n^{\gamma-1/2})$ where ξ_n is evaluated in, and $\gamma < 1/2$. Define

$$\xi_{sn} = -n^{\gamma-1/2} + 2\, n^{\gamma-1/2}\, 3^{-m_n}\, s, \quad s = 0, 1, .., 3^{m_n},$$

where 3^{m_n} is of the same order than $n^{-(\gamma-1/2)/4}$. Note that the greed becomes thinner as the sample size grows and the interval also decreases with n. Let ξ_n^{opt} be the value optimizing (18). For such greed, there exits a j such that $0 \le \xi_{jn} - \xi_n^{opt} \le 2\, n^{\gamma-1/2}\, 3^{-m_n}$

Define for $i = 1, 2, .., n$

$$
\begin{aligned}
X_{is}^{l} &= X_i(1 - 2\, n^{\gamma-1/2}\, 3^{-m_n} \xi_{sn}^{-1} 1(X_i > 0)), \\
X_{is}^{u} &= X_i(1 - 2\, n^{\gamma-1/2}\, 3^{-m_n} \xi_{sn}^{-1} 1(X_i < 0)).
\end{aligned}
$$

This variables satisfy the relationship (see Mukantseva, 1977 and Boldin, 1982), $X_{ij}^{l}\xi_{jn} \leq X_i \xi_n^{opt} \leq X_{ij}^{u}\xi_{jn}$. Suppose for the moment that $h_i(t) \geq 0$ $a.s.$. Then, from the inequalities above we can write

$$
\begin{aligned}
& \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ 1\left(\varepsilon_i \leq X_i \xi_n^{opt}\right) - F(X_i \xi_n^{opt}) - 1\left(\varepsilon_i \leq 0\right) + F(0) \right\} h_i(t) \\
\geq\ & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ 1\left(\varepsilon_i \leq X_{ij}^{l}\xi_{jn}\right) \} - F(X_{ij}^{l}\xi_{jn}) - 1\left(\varepsilon_i \leq 0\right) + F(0) \right\} h_i(t) \\
& - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ F(X_{ij}^{u}\xi_{jn}) - F(X_{ij}^{l}\xi_{jn}) \right\} h_i(t) \\
=\ & b_{1n}(t) - b_{2n}(t)
\end{aligned}
$$

Analogously

$$
\begin{aligned}
& \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ 1\left(\varepsilon_i \leq X_i \xi_n^{opt}\right) - F(X_i \xi_n^{opt}) - 1\left(\varepsilon_i \leq 0\right) + F(0) \right\} h_i(t) \\
\leq\ & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ 1\left(\varepsilon_i \leq X_{ij}^{u}\xi_{jn}\right) - F(X_{ij}^{u}\xi_{jn}) - 1\left(\varepsilon_i \leq 0\right) + F(0) \right\} h_i(t) \\
& + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ F(X_{ij}^{u}\xi_{jn}') - F(X_{ij}^{l}\xi_{jn}) \right\} h_i(t) \\
=\ & b_{3n}(t) + b_{2n}(t)
\end{aligned}
$$

Notice that there is not loss of generality assuming $h_i(t) \geq 0$, because if it does not hold, we can always write $h_i(t) = |h_i(t)| \left(1(h_i(t) \geq 0) - 1(h_i(t) < 0)\right)$, and dividing the left hand side of (18) in two terms, everything follows using some arguments as for $h_i(t) \geq 0$.

Therefore, it suffices to show that $b_{kn}(t)$ vanish for $k = 1, 2, 3$. Applying Liptschitz condition

$$
b_{2n}(t) \;=\; K \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_{ij}^{u}\xi_{jn} - X_{ij}^{l}\xi_{jn}) h_i(t)
$$

$$
\begin{aligned}
&= K\xi_{jn}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(X_{ij}^{u}-X_{ij}^{l}\right)h_i(t)\\
&= K\xi_{jn}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i 2\, n^{\gamma-1/2}\, 3^{-m_n}\xi_{jn}^{-1}\left(1(X_i>0)-1(X_i<0)\right)h_i(t)\\
&= K\left(2\, n^{\gamma}\, 3^{-m_n}\right)\frac{1}{n}\sum_{i=1}^{n}|X_i|\, h_i(t)\\
&= o_p(1). \qquad (19)
\end{aligned}
$$

The limit in (19) has been calculated under **A.1** and **A.2**, and because $n^{\gamma}\, 3^{-m_n} \sim n^{\gamma}\, n^{(\gamma-1/2)/4} = n^{(5\gamma-1/2)/4}$ and it vanishes for any $\gamma < 1/10$. Now

$$
\begin{aligned}
b_{3n}(x) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(1\{\varepsilon_i \le X_{ij}^{u}\xi_{jn}\} - F(X_{ij}^{u}\xi_{jn})\right)h_i(t)\\
&\quad -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(1\{\varepsilon_i \le 0\} - F(0)\right)h_i(t)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
E(b_{3n}(t)^2) &\le E\left\{E\left\{1\left(|\varepsilon_1| \le \left|X_{1j}^{u}\xi_{jn}\right|\right) - \Pr\left(|\varepsilon_1| \le \left|X_{1j}^{u}\xi_{jn}\right|\right) \mid X_1\right\}^2 h_i(t)\right\}\\
&\to 0 \text{ as } n\to\infty.
\end{aligned}
$$

Finally, the convergence of $b_{1n}(x)$ follows using the same arguments. ■

Lemma 2

Under the conditions of the previous Lemma, uniformily in x

$$
\lim_{n\to\infty}\Pr\left(\sup_{\|\xi_n\|\le n^{\gamma-1/2}}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(1\{\varepsilon_i \le Z_i'\xi_n\} - 1\{\varepsilon_i \le 0\} - f(0)Z_i'\xi_n\right)h_i(t)\right| > \delta\right) = 0.
$$

Proof. Note that by the Mean Value Theorem

$$
\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{F(Z_i'\xi_n)-F(0)\right\}h_i(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}f(Z_i'\bar{\xi}_n)Z_i'\xi_n h_i(t) \qquad (20)\\
&= f(0)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_i'\xi_n h_i(t) + o_p(1). \qquad (21)
\end{aligned}
$$

where $\bar{\xi}_n \in (0,\xi_n)$. Now the result follows substituting (20) in Lemma 1. ■

LEMMATA B

The Lemmas below are applied to prove the results in Section 3.

Lemma 3

Under the conditions of Proposition (4), with probability one and uniformily in x

$$\lim_{n\to\infty} \Pr{}^* \left\{ \sup_{|\xi_n|\leq n^{\gamma-1/2}} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{1\,(\varepsilon_i^* \leq Z_i'\xi_n) - F\,(Z_i'\xi_n) - 1\,(\varepsilon_i^* \leq 0) + F(0)\}\, h_i(t) \right.\right.$$

$$\left.\left. +d_n(x,\xi_n) \right| > \delta \right\} = 0,$$

where $d_n(t) = o_p(1)$ is constant with respect to $\Pr{}^*$.

Proof.

Proceed like in the proof of Lemma 1. Everything works in the same way except the convergence to 0 of the analogous expressions $b_{2n}^*(t)$ and $b_{3n}^*(t)$. Taking into account that conditionally on the sample, the regressors are regarded as fixed,

$$\begin{aligned} b_{2n}^*(t) &= K\left(2n^{\gamma}3^{-m_n}\right)\frac{1}{n}\sum_{i=1}^{n}|X_i|\;\, h_i(t) \quad a.s. \\ &= o(1) \quad a.s. \end{aligned}$$

To prove negligibility of $b_{3n}^*(x)$ notice that

$$\begin{aligned} b_{3n}^*(t) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{1\left(\varepsilon_i^* \leq X_{ij}^u\xi_{jn}\right) - F_n(X_{ij}^u\xi_{jn}) - 1\,(\varepsilon_i^* \leq 0) + F_n(0)\right\} h_i(t) \\ &\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{F_n(X_{ij}^u\xi_{jn}) - F(X_{ij}^u\xi_{jn}) - F_n(0) + F(0)\right\} h_i(t) \\ &= +d_n(x,\xi_{jn}) \end{aligned}$$

and $b_{4n}^*(t) \xrightarrow{2} 0$ $a.s.$ is proved analogously as in the proof of Lemma 2. ■

Lemma 4

Under the conditions of the previous Lemma, with probability one

$$\lim_{n\to\infty} \Pr{}^* \left(\left| \sup_{|\xi_n|\leq n^{\gamma-1/2}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (1\{\varepsilon_i^* \leq Z_i'\xi_n\} - 1\{\varepsilon_i^* \leq 0\} + f(0) Z_i'\xi_n)\, h_i(t) \right.\right.$$

$$+ d_n(x,\xi_n)\Big| > \delta\Big) = 0.$$

Proof.

Note that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(F(Z_i'\xi_n) - F(0)\right) h_i(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} f(Z_i'\bar{\xi}_n)Z_i'\xi_n h_i(t) \quad (22)$$

$$= f(0)\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i'\xi_n h_i(t) + o(1), \quad (23)$$

because the regressors are conditionally fixed, where $\bar{\xi}_n \in (0,\xi_n)$. Now apply (22) to Lemma 3 and the proof is completed. ■

Lemma 5.

Suppose the bootstrap random variable W_n^* verifies the decomposition $W_n^* = W_{1n}^* + W_{2n}$ where W_{1n}^* depends on the bootstrap sample and W_{2n} depends only on the original sample (it is constant conditionally on the drawn sample). If $W_{1n}^* \xrightarrow{d} W$ with probability 1 and $W_{2n} \xrightarrow{p} 0$. Then $W_n^* \xrightarrow{d} W$ in probability.

Proof. Let G_n^*, F_n^* and F be the distribution function of W_n^*, W_{1n}^* and W respectively. Let $d(\cdot,\cdot)$ a metric on $\mathcal{F}$, the space of the distribution function and suppose that $d(\cdot,\cdot)$ metricize the weak topology. That is, if U_n and U are two random variables with distribution function H_n, and H respectively,

$$U_n \xrightarrow{d} U \Leftrightarrow \lim_{n\to\infty} d(H_n, H) = 0$$

We will prove the validity of bootstrap approximation using the equivalence

$$W_n^* \xrightarrow{d} W \text{ in probability} \Leftrightarrow d(G_n^*, F) \xrightarrow{p} 0. \quad (24)$$

It is known that

$$\bar{d}(F,H) = \sup_{h\in BL}\left|\int h(x)\,dF - \int h(x)\,dH\right|$$

$$BL = \left\{h : |h(x) - h(y)| \le |x-y|\,; \sup_x |h(x)| \le 1\right\}$$

metricize the weak topology (see Dudley, 1989, th. 11.3.2). The functions in BL are Bounded Lipschitz.

We have by hypothesis that $W_n^* = W_{1n}^* + W_{2n}$. Because W_{2n} is constant conditionally on $\mathcal{X}$, we can write

$$G_n^*(x) = \Pr\left(W_n^* \leq x \mid \mathcal{X}\right) = F_n^*(x - W_{2n})$$

We will prove that $\bar{d}(G_n^*, F) \xrightarrow{p} 0$.

$$\begin{aligned}
\bar{d}(G_n^*, F) &= \sup_{h \in BL} \left| \int h(x)\, dF_n^*(x - W_{2n}) - \int h(x)\, dF(x) \right| \\
&= \sup_{h \in BL} \left| \int h(x + W_{2n})\, dF_n^*(x) - \int h(x)\, dF(x) \right| \\
&\leq \sup_{h \in BL} \left| \int h(x + W_{2n})\, dF_n^*(x) - \int h(x)\, dF_n^*(x) \right| \\
&\quad + \sup_{h \in BL} \left| \int h(x)\, dF_n^*(x) - \int h(x)\, dF(x) \right| \\
&= a_{1n}(x) + a_{2n}(x).
\end{aligned}$$

because of (24), $a_{2n}(x) \xrightarrow{p} 0$. Now applying the properties of BL

$$\begin{aligned}
a_{1n}(x) &= \sup_{h \in BL} \left| \int h(x + W_{2n}) - h(x)\, dF_n^*(x) \right| \\
&\leq \sup_{h \in BL} \int |W_{2n}|\, dF_n^*(x) \\
&= |Z_n| \xrightarrow{p} 0.
\end{aligned}$$

and the proof is completed. ■

Lemma 6.

Suppose the bootstrap random variable W_n^* verify the decomposition $W_n^* = W_{1n}^* W_{2n}$ where W_{1n}^* depends on the bootstrap sample and W_{2n} depends only on the original data (it is constant conditionally on the drawn sample). If $W_{1n}^* \xrightarrow{d} W$ with probability 1, and $W_{2n} \xrightarrow{a.s.} 1$. Then $W_n^* \xrightarrow{d} W$ with probability 1.

Proof. Using the same notation that in the proof of Lemma 5,

$$W_n^* \xrightarrow{d} W \text{ with probability } 1 \Leftrightarrow d(G_n^*, F) \xrightarrow{a.s.} 0. \tag{25}$$

We have by hypothesis that $W_n^* = W_{1n}^* W_{2n}$. Because W_{2n} is constant conditionally on $\mathcal{X}$, we can write

$$G_n^*(x) = \Pr\left(W_n^* \leq x \mid \mathcal{X}\right) = F_n^*\left(\frac{x}{W_{2n}}\right)$$

We will prove that $\bar{d}(G_n^*, F) \stackrel{a.s.}{\to} 0$.

$$\begin{aligned}
\bar{d}(G_n^*, F) &= \sup_{h \in BL} \left| \int h(x)\, dF_n^*(x/W_{2n}) - \int h(x)\, dF(x) \right| \\
&= \sup_{h \in BL} \left| \int h(xW_{2n})\, dF_n^*(x) - \int h(x)\, dF(x) \right| \\
&\leq \sup_{h \in BL} \left| \int h(xW_{2n})\, dF_n^*(x) - \int h(x)\, dF_n^*(x) \right| \\
&\quad + \sup_{h \in BL} \left| \int h(x)\, dF_n^*(x) - \int h(x)\, dF(x) \right| \\
&= a_{1n}(x) + a_{2n}(x).
\end{aligned}$$

because of (25) $a_{2n}(x) \stackrel{a.s.}{\to} 0$. Now applying the properties of BL

$$\begin{aligned}
a_{1n}(x) &= \sup_{h \in BL} \left| \int h(xW_{2n}) - h(x)\, dF_n^*(x) \right| \\
&\leq \sup_{h \in BL} \int |h(xW_{2n}) - h(x)|\ dF_n^*(x) \\
&\leq \sup_{h \in BL} \int |(W_{2n} - 1)\, x|\ dF_n^*(x) \\
&= |W_{2n} - 1| \int |x|\ dF_n^*(x) \stackrel{a.s.}{\to} 0,
\end{aligned}$$

because $W_{2n} \stackrel{a.s.}{\to} 1$, and the proof is completed. ■

REFERENCES

[1] Amemiya, T., (1982): "Two stage least absolute deviations estimators," *Econometrica*, 50, pp. 689-711.

[2] Andrews, D.W.K., (1996): "A Conditional Kolmogorov Test," *Econometrica*, forthcoming.

[3] Bierens, H. (1982): "Consistent Model Specification Test," *Journal of Econometrics,* 20, pp. 105-134.

[4] Bierens, H. (1990): "A Consistent Conditional Moment Test of Functional Form," *Econometrica,* 58, pp. 1443-1458.

[5] Bierens, H. and W. Ploberger (1997): "Asymptotic theory of integrated conditional moment test," Econometrica, forthcoming.

[6] Bloomfield, P. and W.L. Stieger (1983): *Least Absolute Deviations. Theory, Applications and Algorithms.* P. Huber and M. Rosenblatt Eds., Boston: Birkhäuser.

[7] Boldin, M.V. (1982): "Estimation of the distribution of noise in a autoregression scheme," *Theory of Probability and its Applications*, 27, pp 866-871.

[8] Buchinsky, M. (1994): "Changes in the U.S. wage structure 1963-1987: Application of Quantile Regression," *Econometrica*, 72, pp. 405-458.

[9] Buchinsky, M. (1995): "Quantile regression, Box-Cox transformation model, and the U.S. wage structure, 1963-1987," *Journal of Econometrics*, 65, pp. 109-154.

[10] Chamberlain, G. (1991): "Quantile regression, Censoring and the Structure of Wages," Discussion paper 1558, Harvard Institute of Economic Research.

[11] De Angelis, D., P. Hall, and G.A. Young (1993): "Analytical and Bootstrap Approximations to Estimator Distributions in L^1 Regression," *Journal of the American Statistical Association,* 88, pp. 1310-1316.

[12] De Jong R. M. and H.J. Bierens (1994): "On limit behavior of a Chi-Square type test if the number of conditional moments tested approaches infinity," *Econometric Theory*, 9, pp. 70-90.

[13] De Jong R. M. (1996): "The Bierens test under data dependence," *Journal of Econometrics*, 72, pp. 1-32.

[14] Delgado, M. (1993): "Testing the equality of nonparametric regression curves," *Statistics & Probability Letters,* 17, pp. 199-204.

[15] Diebolt, J. (1995): "A nonparametric test for the regression function: Asymptotic theory," *Journal of Statistical Planning and Inference*, 44, pp. 1-17.

[16] Dudley, R. M. (1989): *Real Analysis and Probability.* London: Chapman and Hall.

[17] Gaenssler, P. and W. Stute (1979): "Empirical processes: A Survey of results for independent and identically distributed random variables," *Annals of Probability,* 7, pp. 193-243.

[18] Granger, C.W.J., H. White and M. Kamstra (1989): "Interval Forecasting . An Analysis based upon ARCH-quantil Estimators," *Journal of Econometrics*, 40, pp. 87-96.

[19] Hahn, J. (1995): "Bootstrapping quantile regression estimators," *Econometric Theory*, 11, pp. 105-121.

[20] Hong-zhy, A and C. Bing (1991): "A Kolmogorov-Smirnov type statistic with application to test for nonlinearity in time series," *International Statistical Review,* 59, pp. 287-307.

[21] Koenker R. and G. Basset (1978): "Regression Quantiles," *Econometrica*, 46, pp.33-50.

[22] Koenker R. and G. Basset (1982): "Robust test for heteroscedasticity based on regression quantiles," *Econometrica*, 50, pp. 43-61.

[23] Koenker, R. and V. D'Orey (1987): "Computing Regression Quantiles," *Journal of the Royal Statisitcal Society, Serie C,* 36 pp. 383-393.

[24] Mukantseva, L.A. (1977): "Testing normality in one-dimensional an multidimensional linear regressions," *Theory of Probabability and its Applications*, 22, pp. 591-602.

[25] Oberhoffer, (1982): "The consistency of nonlinear regression minimizing the L_1-norm," Annals of Statistics, 12, pp. 316-319.

[26] Phillips, P.C.B. (1991): "A shortcut to LAD estimator asymptotics," *Econometric Theory*, 7, pp. 450-463.

[27] Pollard, D. (1991): "Asymptotics for Least Absolute Deviations Regression estimators," *Econometric Theory*, 7, pp. 186-199.

[28] Powell, J.L. (1984): "Least absolute deviations estimation for the censored regression model," *Journal of Econometrics*, 25, pp. 303-325.

[29] Powell, J.L. (1986): "Censored regression quantiles," *Journal of Econometrics*, 32, pp. 143-155.

[30] Ruppert, D. and R.J. Carroll. (1980): "Trimmed Least Squares Estimation in the Lineal Model," *Journal of the American Statistical Association*, 75, pp. 828-838.

[31] Stute, W. (1995): "Nonparametric Model Checks for Regression," *Annals of Statistics,* forthcoming.

[32] Stute, W., W.G. Manteiga and M. Presedo (1996) "Bootstrap Approximations in Model Checks for Regression," *Journal of American Statistics Association*, forthcoming.

[33] Su, J.Q. and L.J. Wei (1991): "A lack of fit test for the mean function in a generalized linear model," *Journal of the American Statistical Association,* 86, pp. 420-426.

[34] Weiss A. (1991): "Estimating nonlinear dynamic models using least absolute error estimation". *Econometric Theory*, 7, pp. 46-58.

TABLE I

Asymptotic Test.

$H_0 : Q_{0.5}(Y \mid X) = \theta_{01}$

True Model: $Q_{0.5}(Y \mid X) = 4 + \beta_{02}X$.

		Normal			Lognormal		
β_{02}	α	n=30	n=50	n=100	n=30	n=50	n=100
$\beta_{02} = 0$	0.10	0.140	0.106	0.099	0.140	0.106	0.099
	0.05	0.064	0.059	0.042	0.064	0.059	0.042
	0.01	0.015	0.010	0.013	0.015	0.010	0.013
$\beta_{02} = 1$	0.10	0.189	0.385	0.619	0.619	0.914	0.999
	0.05	0.110	0.265	0.492	0.484	0.864	0.991
	0.01	0.024	0.090	0.276	0.235	0.666	0.956
$\beta_{02} = 2$	0.10	0.528	0.970	0.990	1.000	0.873	1.000
	0.05	0.378	0.922	0.975	0.999	0.790	1.000
	0.01	0.164	0.722	0.903	0.994	0.551	1.000
$\beta_{02} = 5$	0.10	0.989	1.000	1.000	1.000	1.000	1.000
	0.05	0.964	1.000	1.000	1.000	1.000	1.000
	0.01	0.826	1.000	1.000	0.990	1.000	1.000

TABLE II

Bootstrap test based on 500 bootstrap samples.

$H_0 : Q_{0.5}(Y \mid X) = \beta_{01} + \beta_{02}X$

True Model: $Q_{0.5}(Y \mid X) = 4 + \beta_{02}X + \beta_{03}X^2$.

		Normal			Lognormal		
	α	n=30	n=50	n=100	n=30	n=50	n=100
$\beta_{02} = 0$, $\beta_{03} = 0$	0.10	0.100	0.096	0.097	0.103	0.089	0.100
	0.05	0.052	0.050	0.051	0.053	0.046	0.052
	0.01	0.009	0.011	0.010	0.015	0.013	0.010
$\beta_{02} = 5$, $\beta_{03} = 0$	0.10	0.104	0.091	0.100	0.104	0.091	0.100
	0.05	0.055	0.047	0.052	0.055	0.047	0.052
	0.01	0.015	0.013	0.008	0.015	0.013	0.008
$\beta_{02} = 5$, $\beta_{03} = 1$	0.10	0.103	0.114	0.138	0.149	0.199	0.327
	0.05	0.057	0.062	0.087	0.084	0.131	0.232
	0.01	0.015	0.014	0.028	0.024	0.005	0.095
$\beta_{02} = 5$, $\beta_{03} = 5$	0.10	0.214	0.471	0.784	0.593	0.977	1.000
	0.05	0.137	0.349	0.697	0.467	0.947	0.999
	0.01	0.040	0.173	0.466	0.232	0.822	0.997
$\beta_{02} = 5$, $\beta_{03} = 25$	0.10	0.956	1.000	1.000	1.000	1.000	1.000
	0.05	0.912	1.000	1.000	1.000	1.000	1.000
	0.01	0.744	0.999	1.000	0.999	1.000	1.000

TABLE III

Bootstrap test based on 500 bootstrap samples.

$H_0 : Q_{0.5}(Y \mid X) = \beta_{01} + \beta_{02}X + \beta_{03}X^2$

True Model: $Q_{0.5}(Y \mid X) = 4 + 5X + 25X^2 + \beta_{04}X^3$.

		Normal			Lognormal		
	α	n=30	n=50	n=100	n=30	n=50	n=100
$\beta_{04} = 0$	0.10	0.094	0.113	0.096	0.103	0.104	0.101
	0.05	0.048	0.065	0.041	0.047	0.054	0.043
	0.01	0.015	0.020	0.013	0.014	0.016	0.010
$\beta_{04} = 25$	0.10	0.130	0.423	0.857	0.229	0.865	0.999
	0.05	0.062	0.281	0.786	0.131	0.774	0.999
	0.01	0.090	0.104	0.597	0.042	0.550	0.990
$\beta_{04} = 50$	0.10	0.251	0.916	1.000	0.464	0.996	1.000
	0.05	0.145	0.809	1.000	0.291	0.987	1.000
	0.01	0.094	0.543	0.999	0.094	0.962	1.000
$\beta_{04} = 125$	0.10	0.587	1.000	1.000	0.841	1.000	1.000
	0.05	0.353	1.000	1.000	0.668	1.000	1.000
	0.01	0.110	0.999	1.000	0.287	1.000	1.000