Working Paper 98-54
Statistics and Econometrics Series 24
June 1998

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34)91 624-9849

ASYMPTOTIC AND BOOTSTRAP SPECIFICATION TESTS OF NONLINEAR IN VARIABLES ECONOMETRIC MODELS.

Miguel A. Delgado, Manuel A. Domínguez and Pascal Lavergne*

Abstract

We address the issue of consistent specification testing in general econometric models defined by multiple moment conditions. We develop two classes of moment conditions based tests. The first class of tests depends upon nonparametric functions that are estimated by kernel smoothers. The second class of tests depends upon a marked empirical process. Asymptotic and bootstrap versions of these tests are formally justified, and their finite sample performances are investigated by means of Monte-Carlo experiments.

Keywords:
Specification testing; Nonlinear in variable models; Smoothers; Marked empirical processes; Wild 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; Dominguez, Departamento de Estadística y Economtería, Universidad Carlos III de Madrid; Lavergne, INRA-ESR, BP 27, 31326 Castanet-Tolosan Cedex Toulouse, France. This work was done while the third author was visiting Universidad Carlos III, Madrid.Financial support from the European Commission under grant SBR-9631212 and from the Spanish Dirección General de Esneñanza Superior reference PB95-0292 is gratefully acknowledge.

1 INTRODUCTION

A noticeable amount of recent statistical and econometric work has been devoted to consistent specification testing of univariate regression models. Two separate methodologies have been developed in the literature. The first approach consists of comparing parametric (or semiparametric) estimates with semi or nonparametric estimates, as done for instance in Härdle and Mammen (1993), Horowitz and Härdle (1993), Hong and White (1995), Gozalo (1993), Fan and Li (1996), Zheng (1996) and Rodriguez-Campos, González-Manteiga and Cao (1997). The second approach considers tests based on stochastic processes indexed by a "nuisance" parameter taking on an infinite number of values. This testing strategy was first proposed by Bierens (1982, 1990) and considered by Hong-Zhi and Bin (1991), Su and Wei (1991), Delgado (1993), Diebolt (1993), Andrews (1997), Bierens and Ploberger (1997), De Jong and Bierens (1994), Delgado and Domínguez (1997) and Stute (1997) among others.

In this paper, we address the issue of specification testing in general parametric econometric models. Specifically, we consider models defined by multiple conditional moment restrictions that can be nonlinear in endogenous variables. This includes models with nonlinear transformations in the endogeneous variables and nonlinear in variables simultaneous equation models. In this aim, we simultaneously follow the two leading methodologies applied for specification testing of univariate regression models and extend them for jointly testing conditional moment restrictions. As a matter of fact, both approaches can be interpreted in the M-testing framework developed by Newey (1985) and Tauchen (1985), where unconditional restrictions are tested in place of conditional ones. The restrictions we consider are tailored to ensure consistency of the related testing procedures against any alternative. On the one hand, we consider moment conditions depending on unknown functions (regression curves and probability density functions), which can be estimated by means of kernel smoothers or other nonparametric estimation techniques. On the other hand, we consider an infinite number of orthogonality conditions indexed by a "nuisance" parameter. A test statistic can then be built as a functional of the empirical process indexed by this nuisance parameter.

We first study the asymptotic behavior of the proposed test statistics. However, relying only upon asymptotic analysis in implementing those testing procedures presents some difficulties. For the tests based on smoothers, the approximation quality of the null distribution of the test may be poor in small samples and is expected to be sensitive to the choice of a smoothing (or bandwidth) parameter and to the dimension of exogenous variables. For the tests based on empirical processes, the problem is even more acute, as the asymptotic null distribution depends on the (unknown) data generating process and can be easily tabulated only in exceptional circumstances. Therefore, we propose to approximate critical values of each test by "wild" bootstrap procedures, as done in regression contexts by Härdle and Mammen (1993) and Wang and Li (1996) for specification tests based on smoothers, and by Su and Wei (1991) and Stute, González-Manteiga and Presedo (1998) in specification tests based on certain empirical processes. Specifically, we propose two different bootstrap methods. The first one is based on test statistics obtained by plugging-in bootstraped analogs of the residuals in the initial statistics, as proposed in different contexts by Su and Wei (1991), Lewbel (1995), De Jong (1996) and Hansen (1996) among others. The consistency of the resulting bootstrap tests is formally justified and their behavior in small samples is illustrated by means of a Monte-Carlo experiment. The second method is the classical one, namely it uses conditional bootstrap to built bootstrap analogs to the sample test statistic. However, the formal asymptotic analysis of the resulting tests is complicated by the fact that the model can be nonlinear in the endogenous variables, so that a general asymptotic theory requires very specific conditions which are not needed when the model is linear in endogenous variables. We therefore discuss these difficulties and provide evidences on the performances of both bootstrap methods in small samples.

The paper is organized as follows. In the next section, we present the tests statistics and discuss their asymptotic properties. In Section 3, we study the bootstrap tests approximations, we justify the asymptotic validity of the bootstrap tests based on asymptotic expansions and we finally discuss implementation of the classical bootstrap tests. The small sample performances of the proposed asymptotic and bootstrap tests is studied in Section 4 by means of a Monte Carlo experiment. Proofs are confined to the Appendix.

2 ASYMPTOTIC TESTS

Suppose that $\mathcal{Y}_n = \{(Y_i, X_i), i = 1, ..., n\}$ is a sample of independent observations, identically distributed as the vector (Y, X), where $X = \left(X^{(1)}, X^{(2)}, ..., X^{(q)}\right)'$ takes values in $\mathbb{R}^q$ and Y takes values in $\mathbb{R}^d$. The variables are supposed to be related according to a postulated parametric model which specification is determined by conditional moment restrictions of the form

$$H_0 : \exists \theta_0 \in \Theta : \Pr\{E[\psi(Y, X; \theta_0) \mid X] = 0\} = 1.$$

Here θ_0 is an unknown vector of parameters, $\Theta \subset \mathbb{R}^p$ is the parameter space, $\psi : \mathbb{R}^d \times \mathbb{R}^q \times \mathbb{R}^p \to \mathbb{R}^m$ is a vector of known functions such that $E\|\psi(Y, X; \theta_0)\| < \infty$, where $\|\cdot\|$ denotes Euclidean norm. The alternative hypothesis H_1 of incorrect specification is the negation of H_0. Most econometric models can be written in this form. It includes nonlinear regression models, but also general models that involve nonlinear transformations in the endogenous variables (e.g. Box-Cox transform) and nonlinear simultaneous equations models.

It turns out that H_0 is equivalent to any of the following conditions:

$$\exists\, \theta_0 \in \Theta : E\left[\psi(Y, X; \theta_0)' W_1(X)\right] = 0, \tag{1}$$

or

$$\exists\, \theta_0 \in \Theta : E\left[\psi(Y, X; \theta_0) W_2(X, x)\right] = 0,\ \forall x \in \mathcal{X}, \tag{2}$$

where $W_1(X) = E[\psi(Y, X; \theta_0) \mid X] f(X)$, $W_2(X, x) = \prod_{j=1}^{q} 1\left(X^{(j)} \leq x^{(j)}\right)$, $1(A)$ is the indicator function of the event A, $x = \left(x^{(1)}, x^{(2)}, ..., x^{(q)}\right)'$ and $\mathcal{X}$ denotes the support of X. Thus, H_0 can be tested by checking any of the above conditions. Other equivalent formulations of H_0 obviously exist, but they belong to one of the above types, i.e. they rely on an unknown weight function or they specify an infinite number of orthogonality conditions depending on a "nuisance" parameter.

The weight function $W_1(.)$ depends on unknown nonparametric functions, hence implementation of tests motivated by (1) requires nonparametric estimation. Define $\psi(\theta) = \psi(Y, X; \theta)$ and $\psi_i(\theta) = \psi(Y_i, X_i; \theta)$, with k-th coordinate denoted by $\psi^{(k)}(\theta)$ and $\psi_i^{(k)}(\theta)$ respectively. Given a suitable estimator $\hat{\theta}_n$ of θ_0, the expectation in (1) can be estimated by

$$T_n = \frac{1}{n}\sum_{i=1}^{n} \psi_i\left(\hat{\theta}_n\right)' \hat{W}_1(X_i),$$

where

$$\hat{W}_1(X_i) = \frac{1}{(n-1)h^q} \sum_{\substack{j=1 \\ j \neq i}}^{n} \psi_j\left(\hat{\theta}_n\right) K_{ij}$$

is a kernel estimator of $W_1(X_i)$, $h = h(n)$ is a positive bandwidth number,

$$K_{ij} = K\left(\frac{X_i - X_j}{h}\right),$$

and $K(.) : \mathbb{R}^q \to \mathbb{R}$ is a symmetric kernel function. As we will show, a properly rescaled version of T_n can be used as a test statistic for checking the specification of the econometric model.

The weight $W_2(.,.)$ depends on the nuisance parameter x, hence (2) specifies an infinite number of orthogonality conditions, one for each x. The expectation in (2) can be estimated by

$$R_n(x) = \frac{1}{n} \sum_{i=1}^{n} \psi_i\left(\hat{\theta}_n\right) W_2(X_i, x),$$

which is an empirical process of dimension m marked by the "residual" functions $\psi_i\left(\hat{\theta}_n\right)$, $i = 1, .., n$. Some suitable function of the process $R_n(.)$ can then be used for testing H_0. We will consider a Cramer-von Mises type statistic, but a similar theory can be derived for any other statistic based on another norm.

The study of the asymptotic properties of T_n and $R_n(.)$ requires the following usual regularity conditions.

A1 Under H_0, $\hat{\theta}_n$ admits the expansion,

$$\hat{\theta}_n = \theta_0 + \frac{1}{n} \sum_{i=1}^{n} \ell(Y_i, X_i; \theta_0) + o_p\left(n^{-1/2}\right),$$

for some interior point θ_0 of Θ, a compact set in $\mathbb{R}^q$, and some vector-valued function ℓ such that

(i) $E[\ell(Y, X; \theta_0)] = 0$,

(ii) $\mathrm{L}_0 = E\left[\ell(Y, X; \theta_0)\, \ell(Y, X; \theta_0)'\right]$ exists.

A2 For all $k = 1, ..., m$, $\psi^{(k)}(\theta)$ has at least two continuous derivatives in an open neighborhood $N(\theta_0)$ of θ_0 with respect to θ, namely $\dot{\psi}^{(k)}(\theta) = \partial\psi^{(k)}(\theta)/\partial\theta'$ and $\ddot{\psi}^{(k)}(\theta) = \partial^2\psi^{(k)}(\theta)/\partial\theta\partial\theta'$, where $E\left\|\dot{\psi}^{(k)}(\theta_0)\right\| < \infty$, and there exists a function $N(.)$ such that $\sup_{\theta \in \mathcal{N}(\theta_0)} \left\|\ddot{\psi}^{(k)}(\theta)\right\| \leq N(Y, X)$, $k = 1, .., m$, with $E[N(Y, X)] < \infty$.

For instance, the estimator $\hat{\theta}_n$ can be a generalized method of moments (GMM) estimator defined as

$$\hat{\theta}_n = \arg\min_{\theta\in\Theta}\left[\sum_{i=1}^{n} A(X_i)\,\psi_i(\theta)\right]' B_n \left[\sum_{i=1}^{n} A(X_i)\,\psi_i(\theta)\right], \qquad (3)$$

for a suitable $r\times m$ matrix of instruments $A(\cdot)$ and weighting matrix B_n depending on the sample. Hence, under regularity conditions,

$$\ell(Y_i, X_i;\theta_0) = -\left(\mathcal{M}_0'\mathcal{B}_0\mathcal{M}_0\right)^{-1}\mathcal{M}_0'\mathcal{B}_0 A(X_i)\,\psi_i(\theta_0),$$

where

$$\frac{1}{n}\sum_{i=1}^{n} A(X_i)\,\dot{\psi}_i(\theta_0) \overset{a.s.}{\to} \mathcal{M}_0 \text{ and } B_n \overset{a.s.}{\to} \mathcal{B}_0,$$

and $\dot{\psi}_i(\theta) = \left(\dot{\psi}_i^{(1)}(\theta)', \dot{\psi}_i^{(2)}(\theta)', ..., \dot{\psi}_i^{(m)}(\theta)'\right)'$, where $\dot{\psi}_i^{(k)}(\theta) = \partial\psi_i^{(k)}(\theta)/\partial\theta'$, $k = 1, .., m$.

2.1 ASYMPTOTIC TEST BASED ON SMOOTHERS

We need the following minimal assumption on the kernel function.

K The kernel function K is even, bounded, integrable, $\lim_{\|u\|\to\infty}\|u\|^q |K(u)| = 0$, $\int_{\mathbb{R}^q} K(u)\,du > 0$ and $\int_{\mathbb{R}^q}\int_{\mathbb{R}^q} K(u)\,K(u+v)\,du\,dv < \infty$.

Since we are not assuming that $K(.)$ integrates to one, T_n will not estimate the expectation in (1), but its product with $\int K(u)\,du$. This is of no relevance with respect to the testing procedure. However, most kernels used in practice integrate to one. We also assume that certain nonparametric functions belong to a general class, defined through a Lipschitz condition as follows.

Definition 1 G^α, $\alpha > 0$, *is the class of functions* $g(.) : \mathbb{R}^q \to \mathbb{R}$, *satisfying: there exists a* $\rho > 0$ *such that for all* $z \in \mathbb{R}^q$, $\sup_{y\in S_{z\rho}} |g(y) - g(z)| / \|y - z\| \leq G(z)$, *where* $S_{z\rho} = \{y : \|y - z\| \leq \rho\}$, *and* $g(.)$ *and* $G(.)$ *have finite* α*-th moments (or are bounded if* $\alpha = +\infty$*).*

Let $\sigma_{kl}(x) = E\left[\psi^{(k)}\psi^{(l)} \mid X = x\right]$, $\sigma_k^4(x) = E\left[\left(\psi^{(k)}\right)^4 \mid X = x\right]$, $\alpha_k(x) = E\left[\left|\psi^{(k)}\right| \mid X = x\right]$, $\gamma_k(x) = E\left[\dot{\psi}^{(k)} \mid X = x\right]$, $k, l = 1, ..., m$ and $\beta(x) = E[N \mid X = x]$, where the functions $\psi^{(k)}$ are evaluated at θ_0. The next assumption summarizes the smoothness conditions on the different nonparametric functions.

S $f(.) \in G^{\infty}$. For all $k, l = 1, ..., m$, $\sigma_{kl}(.) \in G^4$, each element of $\gamma_k(.)$ is in $G^{8/3}, \sigma_k^4(.)$, $\alpha_k(.)$ and $\beta(.)$ are in G^2 and $E\left[\left\|\dot{\psi}^{(k)}\right\|^{8/3}\right] < \infty$.

Finally, we impose the standard assumptions on the rate of convergence of the bandwidth parameter.

B $\lim_{n\to\infty}\left\{h + (nh^q)^{-1}\right\} = 0$.

To justify the asymptotic test, we first derive an asymptotic linearization of T_n.

Lemma 1 *Under H_0 and Assumptions A1, A2, K, S, B,*

$$T_n = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} h^{-q} \psi_i(\theta_0)' \psi_j(\theta_0) K_{ij} + O_p\left(n^{-1}\right).$$

Since the leading term in the asymptotic linearization of T_n is a degenerate U-statistic, its null distribution of T_n is obtained by Theorem 1 in Hall (1984).

Theorem 1 *Under H_0 and Assumptions A1, A2, K, S, B,*

$$nh^{q/2} T_n \xrightarrow{d} N(0, V),$$

where $V = 2E\left\{\sum_{k=1}^{m}\sum_{l=1}^{m} \sigma_{kl}^2(X) f(X)\right\} \int_{\mathbb{R}^q} K(u)^2 du$.

By analogy with T_n, the variance V can be estimated by

$$\hat{V}_n = \frac{2}{n(n-1)h^q} \sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n} \psi_i^{(k)}\left(\hat{\theta}_n\right) \psi_i^{(l)}\left(\hat{\theta}_n\right) \psi_j^{(k)}\left(\hat{\theta}_n\right) \psi_j^{(l)}\left(\hat{\theta}_n\right) K_{ij}^2.$$

Therefore, an asymptotic one-sided test can be based on the statistic $t_n = nh^{q/2} T_n / \hat{V}_n^{1/2}$, as justified by the following corollary.

Corollary 1 *Under Assumptions A1, A2, K, S and B, $t_n \xrightarrow{d} N(0,1)$ under H_0 and $t_n \xrightarrow{p} +\infty$ under H_1 assuming that A1 and A2 hold with some θ_1 in place of θ_0.*[1]

[1]The assumption of θ_1 being an interior point of Θ is not formally needed if we consider directional derivatives through the interior of Θ in the proofs.

Following the line of Theorem 1's proof, we could similarly show that the test is consistent in the direction of contiguous alternatives of the form

$$H_{1n} : E\left[\psi\left(Y, X; \theta_0\right) \mid X\right] = \frac{g(X)}{n^{1/2}h^{q/4}} \text{ a.s.,}$$

for some $\theta_0 \in \Theta$, where $g(.)$ is a generic non null function. Such alternatives converge to the null slower than the parametric rate $n^{-1/2}$. It is also possible to show that the test is able to detect alternatives of the form

$$H_{2n} : E\left[\psi\left(Y, X; \theta_0\right) \mid X\right] = \eta_n(X) \text{ a.s.,}$$

for some $\theta_0 \in \Theta$, where

$$\eta_n(x) = \alpha_n b\left(\frac{x-c}{\gamma_n}\right),$$

where c is a constant and α_n, γ_n are deterministic sequences converging to zero with the sample size at appropriate rates, $b(.)$ is an integrable and continuously differentiable function up to second order, with band limited with $h = o(\gamma_n)$. The magnitude of the indefinite integral of $\eta_n(x)$ is of order $\alpha_n \gamma_n$, which can be of magnitude $O\left(n^{-1/2}\right)$ or smaller by choosing the bandwidth h, α_n and γ_n in a suitable way. These alternatives were first proposed by Rosenblatt (1975), and later used by Fan (1994) in the context of testing the parametric specification of a density function and by Fan and Li (1997) in the context of testing the goodness-of-fit of a parameterized regression model.

2.2 ASYMPTOTIC TEST BASED ON MARKED EMPIRICAL PROCESSES

To justify an asymptotic test, we first derive an asymptotic linearization of $R_n(.)$ using a standard mean value theorem argument.

Lemma 2 *Under H_0 and Assumptions A1, A2,*

$$R_n(x) = \frac{1}{n}\sum_{i=1}^{n} r_i(x) + o_p\left(n^{-1/2}\right) \text{ a.s.,}$$

uniformly in x, where

$$r_i(x) = \left\{\psi_i(\theta_0)\Delta_i(x) + E\left[\dot{\psi}_i(\theta_0)\Delta_i(x)\right] l\left(Y_i, X_i; \theta_0\right)\right\}$$

with $\Delta_i(x) = W_2(X_i, x) = \prod_{j=1}^{q} 1\left(X_i^{(j)} \leq x^{(j)}\right)$, for $X = \left(X_i^{(1)}, .., X_i^{(q)}\right)'$ and $x = \left(x^{(1)}, .., x^{(q)}\right)'$.

The process $R_n(.)$ is a random element in the Skorokhod space $\times_{l=1}^{m} D(\mathbb{R}^q)$, which is defined as the space of all real functions $h(.)$ continuous from above in $\mathbb{R}^q$, such that for any x and every sequence $\{x_n\}$ approaching to x in some quadrant with corner x, $\lim_{n\to\infty} h(x_n)$ exists. The process is extended to $\times_{l=1}^{m} D[\mathbb{R}^q]$ by defining

$$R_n(-\infty) = 0, \text{ and } R_n(\infty) = \frac{1}{n}\sum_{i=1}^{n} \psi_i\left(\hat{\theta}_n\right),$$

see Stute (1997) for a discussion of convergence in this space. From the expansion of Lemma 2 it is straightforward to show that, for any finite s and arbitrary $(x_1, .., x_s)$, $(\sqrt{n}R_n(x_1), .., \sqrt{n}R_n(x_s))$ converges under H_0 to a gaussian random vector $(R_\infty(x_1), .., R_\infty(x_s))$. By showing the tightness of the process, we obtain that $\sqrt{n}R_n(.)$ converges weakly to the m-valued gaussian process $R(.)$ with q parameters, whose projections on $(x_1, .., x_s)$ are $(R_\infty(x_1), .., R_\infty(x_s))$, i.e. a m-variate gaussian sheet with q parameters.

Theorem 2 *Under H_0 and Assumptions A1, A2,*

$$\sqrt{n}R_n(.) \Rightarrow R_\infty(.) \text{ on } \times_{l=1}^{m} D[\mathbb{R}^q],$$

where $R_\infty(.)$ is a centered m-variate gaussian sheet with q parameters and covariance structure $E\left[r_i(x_1)\, r_i(x_2)'\right]$, $\forall\, x_1, x_2 \in \mathbb{R}^q$.

Thus, an asymptotic specification test can be based on some suitable function of $R_n(.)$. We choose to study a Cramer-von Mises' type test statistic defined as $c_n = n\int R_n'(x)\, R_n(x)\, dF_n(x) = \sum_{i=1}^{n} R_n(X_i)' R_n(X_i)$. Alternatively, we could also follow Bierens (1982) and integrate the process with respect to a measure different than the empirical distribution $F_n(.)$. However, the empirical measure seems a natural and convenient choice. The following corollary provides the asymptotic behavior of c_n.

Corollary 2 *Under Assumptions A1 and A2, $c_n \xrightarrow{d} \int_{\mathbb{R}^q} R_\infty(x)' R_\infty(x)\, dF(x)$ under H_0, where $F(.)$ is the distribution function of X, and $c_n \xrightarrow{p} +\infty$ under H_1 assuming that A1 and A2 hold with some θ_1 in place of θ_0.*[2]

[2] The assumption of θ_1 being an interior point of Θ is not formally needed if we consider directional derivatives through the interior of Θ in the proofs.

The practical applicability of a test based on c_n is hampered by the fact that the limiting distribution of the test statistic under the null hypothesis is case-dependent, namely it depends of the unknown data-generating process, and therefore can not be tabulated. One way of solving this problem is to derive case-independent upperbounds of the asymptotic critical values, as proposed by Bierens and Ploberger (1997). Alternatively, we propose in Section 3 to implement bootstrap tests.

It can be shown that the (infeasible) test based on c_n is consistent in the direction of contiguous alternatives of the form

$$H_{3n} : E\left[\psi\left(Y, X; \theta_0\right) \mid X\right] = \frac{g\left(X\right)}{n^{1/2}} \text{ a.s.,}$$

for some $\theta_0 \in \Theta$, where $g\left(.\right)$ is a nonnull generic function.[3] Bierens and Ploberger (1997) and Stute (1997) show the consistency of some related sempirical process based tests under alternatives of type H_{3n} for regression models. The test based on smoothers cannot detect alternatives like H_{3n}, but conversely the test based on a marked empirical process cannot detect alternatives like H_{2n}. This fact was pointed out by Rosenblatt (1975) and has been recently studied by Fan and Li (1997) comparing Bierens' test and a test based on smoothers in the context of specification testing of regression models.

3 BOOTSTRAP TESTS

The practical implementation of the asymptotic testing procedures introduced in the previous section involves some difficulties. On the one hand, asymptotic tests based on smoothers can be very sensitive to the choice of the smoothing parameter, while the approximation to the asymptotic null distribution can be slow, depending on the number of exogenous variables X in the model, see e.g. Härdle and Mammen (1993). On the other hand, the asymptotic behavior of c_n is not distribution free, so that an asymptotic test can only be implemented in rare circumstances.[4] Therefore, bootstrap tests are sorely needed. A bootstrap test relies on the quantiles of a bootstrap test statistic, that are computed from

[3] For H_{3n} to be an alternative that can be distinguished from H_0 at a parametric rate, the function $g\left(.\right)$ must be orthogonal to $E\left[\psi\left(Y, X; \theta_0\right) \mid X = .\right]$, see Bierens and Ploberger (1997) for a discussion on this point.

[4] This is the case when testing the significance of one explanatory variable in a regression model, see Delgado and Domínguez (1995).

artificial samples generated so as they mimic some features of the initial sample. These quantiles can be approximated as accurately as desired by Monte Carlo.

For testing procedures, it is the distribution under the null hypothesis that is of interest, so that the bootstrap sampling must impose the satisfaction of H_0. In the context of regression specification tests, Härdle and Mammen (1993) and Stute, González-Manteiga and Presedo (1997) have pointed out the necessity of applying a residual based bootstrap in order to resample under the null hypothesis of correct specification. In our general framework, the role of the residuals is played by the ψ_i's evaluated at $\hat{\theta}_n$. Let V_i, $i = 1, ..., n$ be random numbers such that the following assumption holds.

V The V_i's are independently distributed with bounded support such that $E(V_i) = 0$ and $E(V_i^2) = 1$, $i = 1, ...n$.

The bootstrap procedure consists in generating a great number of independent and identically distributed random samples of V_i's satisfying Assumption V to obtain bootstrap samples $\hat{\psi}_i^* = \hat{\psi}_i V_i$, $i = 1, ..., n$, that therefore satisfy the null hypothesis. Each bootstrap sample in turn forms a basis to construct a test statistic and the distribution of the resulting bootstrap statistics can therefore be used to estimate the critical regions of the test. Such a procedure for obtaining artificial samples is known as "wild bootstrap" and was introduced by Wu (1986) in the context of heteroskedastic linear models.

We propose two alternative methods of computing bootstrap test statistics. In the first method, we simply plug in the $\hat{\psi}_i^*$'s in place of the initial ψ_i's to obtain a bootstrap test statistic. This naive *plug-in* procedure is similar to the one employed by Su and Wei (1991), Lewbel (1995), De Jong (1996) and Hansen (1996) in other contexts. In the second method, the bootstrapped $\hat{\psi}_i^*$'s are used to obtain a resample $\mathcal{Y}_n^* = \{(Y_i^*, X_i), i = 1, .., n\}$ from the original sample $\mathcal{Y}_n$. We subsequently compute a test statistic from $\mathcal{Y}_n^*$ in the same way as the original test statistic has been derived from the original sample. This *analog* bootstrap method generalizes the method in Härdle and Mammen (1993), Wang and Li (1997) and Stute, González-Manteiga and Presedo (1997) in specification testing of regression models. The second method requires that we can retrieve the Y_i^*'s from the $\hat{\psi}_i^*$'s, which may not always be possible. Moreover, it is computationally much more demanding than the first one. However, analog bootstrap tests are expected to share the excellent performances enjoyed, in general, by bootstrap tests.

Suppose η_n^* is the bootstrap statistic which distribution, conditional on the sample, is used for approximating the distribution of η_n (i.e. η_n can be c_n or t_n). Suppose that under H_0, $\eta_n \xrightarrow{d} \eta$. First we must show that under H_0, $\eta_n^* \xrightarrow{d^*} \eta$ in probability (a.s.) where $\xrightarrow{d^*}$ in probability (a.s.) means weak convergence in bootstrap distribution in probability (a.s.) according to the following definition.

Definition 2 *Define* $\Pr^*(\cdot) = \Pr(\cdot \mid \mathcal{Y}_n)$. *Let* η_n^* *be a bootstrap test statistic. It is said that* η_n^* *converges weakly (almost surely) in bootstrap distribution to the random variable* η, *with distribution function* $H(.)$, *and it is denoted by* $\eta_n^* \xrightarrow{d^*} \eta$ *in probability (almost surely), whenever the sequence of random variables* $\Pr^*(\eta_n^* \leq z)$ *converges to* $H(z)$ *in probability (almost surely), for every continuity point* z *of* $H(.)$.

Since under H_0, the conditional distribution of η_n^* consistently estimates the distribution of η_n, we can use the quantiles, obtained from the empirical distribution of the Monte Carlo sample of η_n^*, as estimators of the corresponding quantiles of η_n. To ensure consistency of the bootstrap test, the bootstrap statistic η_n^* only needs be bounded in probability (or a.s.) under H_1. For the smooth test, the test statistic is asymptotically pivotal, and hence converges under the alternative hypothesis to the same distribution as under the null hypothesis. For the empirical process based test, the limiting distribution of the test statistic depends upon the data generating process under H_1 as well as under H_0.

3.1 PLUG-IN BOOTSTRAP TESTS

The plug-in version on the smooth test approximates the distribution of t_n by the conditional distribution of the bootstrap statistic

$$t_n^* = \frac{nh^{q/2}T_n^*}{\sqrt{\hat{V}_n}}, \text{ where } T_n^* = \frac{1}{h^q}\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \hat{\psi}_i^{*\prime}\hat{\psi}_j^* K_{ij}.$$

The plug-in version on the empirical process test approximates the distribution of c_n by the conditional distribution of the bootstrap statistic

$$c_n^* = \sum_{i=1}^{n} R_n^*(X_i)' R_n^*(X_i), \text{ where } R_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} r_i^*(x) \text{ and}$$

$$r_i^*(x) = \left\{ \hat{\psi}_i \Delta_i(x) + \left[\frac{1}{n} \sum_{i=1}^{n} \dot{\psi}_i \left(\hat{\theta}_n \right) \Delta_i(x) \right] l\left(Y_i, X_i; \hat{\theta}_n \right) \right\} V_i.$$

The next two theorems formally justify the use of the bootstrap tests.

Theorem 3 *Under V, A1, A2, K, S and B,*

$$t_n^* \xrightarrow{d^*} N(0,1) \text{ in probability}$$

under H_0, and under H_1 assuming that A1 and A2 hold with some θ_1 in place of θ_0.

Theorem 4 *Let $\dot{l}(\theta) = \partial l(Y, X; \theta)/\partial\theta$. Assume that $\sup_{\theta \in \mathcal{N}(\theta_0)} \left\| \dot{l}(\theta) \right\|^2 < \dot{L}(Y, X)$ with $E\left[\dot{L}(Y, X) \right] < \infty$ and that V holds. Then, under H_0 and Assumptions A1, A2,*

$$c_n^* \xrightarrow{d^*} \int_{\mathbb{R}^q} R_\infty(x)' R_\infty(x) \, dF(x) \quad a.s,$$

and under H_1, assuming that A1 and A2 hold with some θ_1 in place of θ_0,

$$c_n^* \xrightarrow{d^*} \int_{\mathbb{R}^q} R_\infty^1(x)' R_\infty^1(x) \, dF(x) \quad a.s.,$$

where $R_\infty^1(.)$ is a gaussian process centered at zero and with the same covariance structure than $R_\infty(.)$ with θ_1 in place of θ_0.

3.2 ANALOG BOOTSTRAP TESTS

When there are as many equations as response variables, i.e., when $m = d$, it can be possible to obtain a resample $\mathcal{Y}_n^* = \{(Y_i^*, X_i), \ i = 1, .., n\}$ from the bootstrapped $\hat{\psi}_i^*$'s. Assume that $\psi(Y, X; \theta) = \psi$ has an unique solution for $Y = Y(\psi, X; \theta)$ in a neighborhood of θ_0 under H_0 and in a neighborhood of θ_1 under H_1. A closed form of the solution is, in general, not available, but numerical methods can be used. Hence we obtain Y_i^* by solving $\psi\left(Y_i^*, X_i; \hat{\theta}_n \right) = \hat{\psi}_i^*$, $i = 1, .., n$. The bootstrap analog of $\hat{\theta}_n$ in (3) is the solution to

$$\hat{\theta}_n^* = \arg\min_{\theta \in \Theta} \phi_n^*(\theta)' B_n \phi_n^*(\theta),$$

where $\phi_n^*(\theta) = n^{-1} \sum_{i=1}^{n} A(X_i) \psi(Y_i^*, X_i; \theta)$. In practice, the extremum estimator $\hat{\theta}_n^*$ may not actually be located, but rather the outcome of finitely many iterations of some iterative procedure. A convenient form of Gauss-Newton iteration towards the solution is

$$\hat{\theta}_n^{*(k+1)} = \hat{\theta}_n^{*(k)} - \left(M_n^*\left(\hat{\theta}_n\right)' B_n M_n^*\left(\hat{\theta}_n\right)\right)^{-1} M_n^*\left(\hat{\theta}_n\right)' B_n \phi_n^*\left(\hat{\theta}_n^{*(k)}\right),$$

for $k = 2, 3, ...,$ where $\hat{\theta}_n^{*(1)} = \hat{\theta}_n$ and $M_n^*(\theta) = n^{-1}\sum_{i=1}^n A(X_i)\dot{\psi}(Y_i^*, X_i; \theta)$.

Let us denote $Z_n^* = o_{p^*}(1)$ a.s. (or in probability) if $\forall\varepsilon > 0, \Pr^*(|Z_n^*| > \varepsilon) \overset{a.s.(p)}{\to} 0$.[5] By construction, we have

$$\sqrt{n}\left(\hat{\theta}_n^{*(2)} - \hat{\theta}_n\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \ell_{in}^*\left(\hat{\theta}_n\right),$$

where

$$\ell_{in}^*(\theta) = -\left(M_n^*(\theta)' B_n M_n^*(\theta)\right)^{-1} M_n^*(\theta) B_n A(X_i)\psi(Y_i^*, X_i; \theta).$$

Therefore, the one-step estimator $\hat{\theta}_n^{*(2)}$ is consistent in bootstrap law, in the sense that $\hat{\theta}_n^{*(2)} = \hat{\theta}_n + o_{p^*}(1)$ a.s., whenever

$$M_n^*\left(\bar{\theta}_n^*\right) = M_n^*\left(\hat{\theta}_n\right) + o_{p^*}(1) \;\; a.s., \; \forall \;\; \bar{\theta}_n^* : \bar{\theta}_n^* = \hat{\theta}_n + o_{p^*}(1) \text{ a.s.}$$

Similarly for all $k \geq 2$, $\hat{\theta}_n^{*(k)}$ is consistent in bootstrap law as

$$\sqrt{n}\left(\hat{\theta}_n^{*(k)} - \hat{\theta}_n\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \ell_{in}^*\left(\hat{\theta}_n\right) + o_{p^*}(1) \text{ a.s.}$$

Let $\psi_i^{**}(\theta) \equiv \left(\psi_i^{**(1)}(\theta)', \psi_i^{**(2)}(\theta)', ..., \psi_i^{**(m)}(\theta)'\right)' = \psi(Y_i^*, X_i; \theta)$ and $\tilde{\theta}_n^* = \hat{\theta}_n^{*(k)}$ for some $k > 1$. The bootstrap analogs of c_n and t_n are thus

$$c_n^{**} = \sum_{i=1}^n R_n^{**}(X_i)' R_n^{**}(X_i), \text{ with } R_n^{**}(x) = \frac{1}{n}\sum_{i=1}^n \psi_i^{**}\left(\tilde{\theta}_n^*\right)\Delta_i(x),$$

and

$$t_n^{**} = \frac{nh^{q/2}T_n^{**}}{\sqrt{\hat{V}_n^*}}, \text{ with } T_n^{**} = \frac{1}{n(n-1)h^q}\sum_{i=1}^n\sum_{\substack{j\neq i\\ j=1}}^n \psi_i^{**}\left(\tilde{\theta}_n^*\right)' \psi_j^{**}\left(\tilde{\theta}_n^*\right) K_{ij},$$

$$\hat{V}_n^* = \frac{2}{n(n-1)h^q}\sum_{s=1}^m\sum_{l=1}^m\sum_{i=1}^n\sum_{\substack{j=1\\ j\neq i}}^n \psi_i^{*(s)}\left(\tilde{\theta}_n^*\right)\psi_i^{*(l)}\left(\tilde{\theta}_n^*\right)\psi_j^{*(s)}\left(\tilde{\theta}_n^*\right)\psi_j^{*(l)}\left(\tilde{\theta}_n^*\right) K_{ij}^2.$$

[5]In this subsection, one can read almost surely (or in probability) any time we use a.s.

For regression models, which are linear in the endogenous variables, the formal asymptotic analysis of bootstrap analogs of the test statistics is almost similar to the one of bootstrap tests based on asymptotic expansions. However, in the general case where the econometric model allows for nonlinear transformation in the endogenous variables, the analysis is much more complicated. One could derive some primitive assumptions that imply validity of our bootstrap tests, similar to those established by Hall and Horowitz (1996). We prefer not to address this issue, but rather we provide some conditions that ensure consistency of the tests based on c_n^{**} and t_n^{**} and discuss the implications for practical implementation of the tests.

Let us first consider the test based on c_n^{**}. Assume that for any $\bar{\theta}_n^*$ (which may depend on the bootstrap sample) such that $\bar{\theta}_n^* = \hat{\theta}_n + o_{p^*}(1)$ a.s., and any measurable function $S(.)$,

$$\frac{1}{n}\sum_{i=1}^{n} S(X_i)\,\dot{\psi}\left(Y_i^*, X_i; \bar{\theta}_n^*\right) = \frac{1}{n}\sum_{i=1}^{n} S(X_i)\,\dot{\psi}\left(Y_i^*, X_i; \hat{\theta}_n\right) + o_{p^*}(1) \text{ a.s.} \tag{4}$$

and

$$E^*\left\{\frac{1}{n}\sum_{i=1}^{n} S(X_i)\,\dot{\psi}\left(Y_i^*, X_i; \hat{\theta}_n\right)\right\} = \frac{1}{n}\sum_{i=1}^{n} S(X_i)\,\dot{\psi}\left(Y_i, X_i; \hat{\theta}_n\right) + o_p(1) \text{ a.s.,} \tag{5}$$

where $E^*(\cdot) = E(\cdot \mid \mathcal{Y}_n)$. Then, it is straightforward to prove that $R_n^{**}(x) = R_n^*(x) + o_{p^*}(1)$ a.s. uniformly en x. The above conditions are therefore sufficient for the validity of the bootstrap test based on c_n^*. Condition (4) can be shown to follow assuming that $\psi(.)$ fulfills conditions similar to Assumption 3 of Hall and Horowitz (1996). Condition (5) is satisfied choosing appropriate V_i's, and the particular parametric form of $\psi(.)$ can help to determine a suitable procedure to generate these random numbers. Moreover, when certain information on the data generating process is available, the V_i's can be chosen in a standard way. For instance, if the conditional density of $\psi(Y, X; \theta)$ given X is symmetric, we can choose V_i such that $\Pr(V_i = 1) = \Pr(V_i = -1) = 1/2$, so that the right-hand side of (5) equals

$$\frac{1}{2n}\sum_{i=1}^{n} S(X_i)\,\dot{\psi}\left(Y\left(\hat{\psi}_i, X_i; \hat{\theta}_n\right), X_i; \hat{\theta}_n\right) + \frac{1}{2n}\sum_{i=1}^{n} S(X_i)\,\dot{\psi}\left(Y\left(-\hat{\psi}_i, X_i; \hat{\theta}_n\right), X_i; \hat{\theta}_n\right),$$

which is a.s. equivalent to $n^{-1}\sum_{i=1}^{n} S(X_i)\,\dot{\psi}\left(Y_i, X_i; \hat{\theta}_n\right)$ under classical regularity conditions.

Different conditions are required for the validity of the test based on t_n^{**}. Assume that for any $\bar{\theta}_n^*$ and $\tilde{\theta}_n^*$ such that $\bar{\theta}_n^* = \hat{\theta}_n + o_{p^*}(1)$ a.s. and $\tilde{\theta}_n^* = \hat{\theta}_n + o_{p^*}(1)$ a.s.,

$$\widehat{S}_{1n}^* = \binom{n}{2}^{-1} \sum_{i<j} \psi_i^*\left(\hat{\theta}_n\right) \dot{\psi}_j^*\left(\bar{\theta}_n^*\right) K_{ij} = O_{p^*}\left(n^{-1/2}\right) \text{ a.s.,} \tag{6}$$

$$\widehat{S}_{2n}^* = \binom{n}{2}^{-1} \sum_{i<j} \dot{\psi}_i^*\left(\bar{\theta}_n^*\right) \dot{\psi}_j^*\left(\tilde{\theta}_n^*\right) K_{ij} = O_{p^*}\left(n^{-1/2}\right) \text{ a.s.,} \tag{7}$$

$$\hat{V}_n^* = \hat{V}_n + o_{p^*}(1) \text{ a.s.} \tag{8}$$

It is then easy to prove under the other regularity assumptions of Theorem 3 that $T_n^{**} = T_n^* + o_{p^*}(1)$ a.s. Following the proof of Theorem 3, (6) holds if

$$\hat{S}_{1n}^* = \hat{S}_{1n} + o_{p^*}(1) \text{ a.s., where } \hat{S}_{1n} = \binom{n}{2}^{-1} \sum_{i<j} \psi_i\left(\hat{\theta}_n\right) \dot{\psi}_i\left(\hat{\theta}_n\right) K_{ij},$$

and (7) holds under a similar condition. Again, if the conditional density of $\psi(Y, X; \theta)$ given X is symmetric, choosing the distribution of the V_i's that assigns equal probability to 1 and -1 ensures (6)-(8).

4 MONTE CARLO STUDY

To investigate the small sample behavior of the tests we have proposed, we have performed Monte Carlo simulations for a well-studied model with a nonlinear transformation in the endogeneous variable, namely the arcsinh transformation. This model was proposed by Burbidge, Magee and Robb (1988) as an alternative to the Box-Cox transformation, and was used by Robinson (1991) among others. The hypothesis of interest here writes

$$H_0 : \exists\,(\lambda_0, \alpha_0, \beta_0) : \Pr\left\{E\left[\frac{\operatorname{arcsin h}(\lambda_0 Y)}{\lambda_0} - \alpha_0 - \beta_0 X \mid X\right] = 0\right\} = 1.$$

We consider the design

$$\frac{\operatorname{arcsin h}(2Y_i)}{2} = 1 + 2X_i + \sin(\delta\pi X_i) + u_i,\; i = 1, \ldots, n,$$

where the u_i's are iids $N(0, 0.5)$ and independently distributed of the X_i's which are iids $N(0,1)$. The null hypothesis DGP_0 corresponds to $\delta = 0$. We investigate three alternatives DGP_1, DGP_2 and DGP_3 corresponding to $\delta = 1, 2$ and 3 respectively. Increasing δ, we obtain higher frequency alternatives that are more difficult to distinguish from pure noise.

The parameters are estimated by GMM with objective function as in (3) and vector of instruments $A(X_i) = (1, X_i, X_i^2)'$ and $B_n = I_3$. For the tests based on smoothers, we choose the bandwidth following the rule-of-thumb, i.e. as $h = dn^{-1/5}$ for different values of d. Three different sample sizes, $n = 50, 100$ and 250, are considered. Tables 1 to 4 report results obtained for each of the considered data generating processes using the five tests proposed in this paper. In each cell are reported the empirical frequencies of rejections of the null hypothesis at a 10% (first row) and a 5% (second row) nominal level.

The empirical sizes of our tests are reported in Table 1. The tests based on t_n and t_n^* are undersized for large and moderate bandwidths, i.e. for d between 2 and 0.5, irrespective of the sample size. Better empirical sizes are obtained for smaller bandwidths. On the whole, the smooth tests based on t_n and t_n^* lead to very similar results. Thus the plug-in bootstrap test does not seem able to improve on the size performances of the asymptotic smooth test for small and moderate samples. Similarly, the test based on c_n^* is undersized in small samples, while its performance improves when the sample size grows. In contrast, the bootstrap analogs of the two tests exhibits empirical sizes very close to nominal sizes. This is attained even for a sample size as small as 50, and for the smooth test irrespective of the bandwidth choice.

We now turn to the study of the empirical power. Under the first alternative DGP_1 (Table 2), the three smooth tests exhibits very good performances. For a sample size of 50, empirical power is lower for very small bandwidths. When the sample size is 100 or higher the empirical power is greater than 85% for any bandwidth and attains 95% in most cases. In contrast, the test based on c_n^* has quite small power against DGP_1 for $n = 50$ or 100. This is largely corrected by the bootstrap analog version of the test, whose empirical power increases with the sample size.

Under the second alternative which has higher frequency than DGP_1, the power of the three smooth tests deteriorates for small samples and large bandwidths. Indeed oversmoothing makes it difficult to detect highly variable periodic alternatives and bootstrap methods are not suitable to correct for this effect. However, the performance of all three tests highly improve for a moderate sample size of 250. The comparison of Tables 1 and 2 indicates that the empirical power of the test based on c_n^{**} is also sensitive to the frequency of the alternative.

Results under alternative DGP_3 (Table 3) confirm these first findings. For bandwidths constants greater than 1.5, the three smooth tests perform poorly even for a sample size of 250. However, their empirical power is acceptable in small samples for a bandwidth constant less than 1 and close to 1 for a sample size of 250. The empirical power curve of these tests exhibits an inverse U-shape with respect to the smoothing number, with a maximum attained by undersmoothing relative to the rule-of-thumb. This fact has already be noted for other smooth specification tests, see Hart (1996). The test based on c_n^* has low power for samples of size 50 and 100. This is partly corrected by the bootstrap analog version of the empirical process based test.

To sum up, our results call for some caution in using plug-in bootstrap tests. For the smooth test, large undersmoothing is required to obtain good size and power properties. For the empirical process test, moderate sample sizes are necessary. In contrast, the bootstrap analogs of these tests appear to enjoy much better properties, though they are computationally more demanding. The bootstrap analog of the smooth test is able to correct for too small empirical sizes of the asymptotic test, and this for a large range of bandwidths. Therefore, only power considerations should drive the choice of the bandwidth parameter for the analog bootstrap tests. For alternatives of varying frequencies, slight undersmoothing with respect to the usual rule-of-thumb seems on the whole to lead to better results. Similarly, among the empirical process based tests, the analog bootstrap form appears to enjoy an accurate size and good empirical power under alternatives with different frequencies.

APPENDIX

Henceforth, $\psi_i^{(k)} = \psi_i^{(k)}(\theta_0)$, $\hat{\psi}_i^{(k)} = \psi_i^{(k)}\left(\hat{\theta}_n\right)$, $\dot{\psi}_i^{(k)} = \dot{\psi}_i^{(k)}(\theta_0)$.

PROOF OF LEMMA 1

Notice that $T_n = \sum_{k=1}^m T_n^{(k)}$, where

$$T_n^{(k)} = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} H_n^{(k)}(Z_i, Z_j) + 2V_{1n}^{(k)} + V_{2n}^{(k)},$$

and $H_n^{(k)}(Z_i, Z_j) = h^{-q} \psi_i^{(k)} \psi_j^{(k)} K_{ij}$,

$$V_{1n}^{(k)} = \frac{1}{n(n-1)h^q} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \psi_i^{(k)} \left(\hat{\psi}_j^{(k)} - \psi_j^{(k)}\right) K_{ij},$$

$$V_{2n}^{(k)} = \frac{1}{n(n-1)h^q} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left(\hat{\psi}_i^{(k)} - \psi_i^{(k)}\right) \left(\hat{\psi}_j^{(k)} - \psi_j^{(k)}\right) K_{ij}.$$

Hence, it suffices to prove that $V_{1n}^{(k)} = O_p\left(n^{-1}\right)$ and $V_{2n}^{(k)} = O_p\left(n^{-1}\right)$, for all $k = 1, ..., m$.

Using a mean value theorem argument,

$$V_{1n}^{(k)} = \left(\hat{\theta}_n - \theta\right)' S_{1n} + \left(\hat{\theta}_n - \theta_0\right)' S_{2n} \left(\hat{\theta}_n - \theta_0\right),$$

where

$$S_{1n} = \frac{1}{n(n-1)h^q} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \psi_i^{(k)} \dot{\psi}_j^{(k)} K_{ij},$$

$$S_{2n} = \frac{1}{n(n-1)h^q} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \psi_i^{(k)} \ddot{\psi}_j^{(k)} \left(\bar{\theta}_n\right) K_{ij}, \ \left\|\bar{\theta}_n - \theta_0\right\| \leq \left\|\hat{\theta}_n - \theta_0\right\|.$$

The order of S_{1n} is obtained applying Lemma 3.1 of Powell, Stock and Stocker (1989), which is reported below.

Lemma 3 *Consider a U-statistic of the form* $U_n = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} H_n(Z_i, Z_j)$, *where* $H_n(Z_i, Z_j)$ *is a symmetric kernel and the* Z_i*s are iids observations. Let* $q_n(Z_i) = E\left[H_n(Z_i, Z_j) \mid Z_i\right]$, $\bar{q}_n = E\left(q_n(Z_i)\right)$ *and* $\hat{U}_n = \bar{q}_n + 2n^{-1} \sum_{i=1}^{n} \left[q_n(Z_i) - \bar{q}_n\right]$.

If $E\left(\left\|H_n(Z_1, Z_2)\right\|^2\right) = o(n)$, *then* $U_n = \hat{U}_n + o_p\left(n^{-1/2}\right)$.

Since S_{1n} is a U-statistic with symmetric kernel

$$H_n(Z_1, Z_2) = \left(h^{-q}\psi_1^{(k)}\dot{\psi}_2^{(k)}K_{12} + h^{-q}\psi_2^{(k)}\dot{\psi}_1^{(k)}K_{12}\right)/2,$$

we have

$$\begin{aligned}
E\left(\|H_n(Z_1, Z_2)\|^2\right) &= \frac{1}{h^{2q}}E\left[\left(\dot{\psi}_1^{(k)}\right)^2\left(\psi_2^{(k)}\right)^2 K_{12}^2\right] \\
&= \frac{1}{h^q}E\left[\left(\dot{\psi}_1^{(k)}\right)^2 \int \sigma_{kk}(X_1 + hu) f(X_1 + hu) K(u)\ ^2du\right] \\
&= \frac{1}{h^q}E\left[\left(\dot{\psi}_1^{(k)}\right)^2 \sigma_{kk}(X_1)\right]\int K(u)^2\ du + o\left(\frac{1}{h^q}\right) \\
&= O\left(\frac{1}{h^q}\right) = o(n)
\end{aligned}$$

by Assumptions B and S, together with Hölder's inequality. As $\bar{q}_n = 0$, we obtain by Lemma 3

$$S_{1n} = \frac{2}{n}\sum_{i=1}^{n} q_n(Z_i) + o_p\left(n^{-1/2}\right),$$

with $q_n(Z_i) = h^{-q}\left(\psi_i^{(k)}E\left[\gamma_k(X_j) K_{ij} \mid X_i\right] + E\left[\psi_j^{(k)}K_{ij} \mid X_i\right]\gamma_k(X_i)\right)/2$. Now it is easy to check that

$$E\left[q_n^2(Z)\right] = E\left[\sigma_{kk}(X)\gamma_k^2(X) f(X)\right]\int K(u)\ du + o(1) = O(1),$$

so that $S_{1n} = O_p\left(n^{-1/2}\right)$ by applying a central limit theorem argument. The order of S_{2n} is obtained using Assumptions $A2$ and S and applying standard kernel manipulations as follows:

$$\begin{aligned}
E\left|S_{2n}\right| &\leq \frac{1}{h^q}E\left\{\left|\psi_1^{(k)}\right| N(Y_2, X_2)\left|K_{12}\right|\right\} \\
&= \frac{1}{h^q}E\left\{N(Y_2, X_2) E\left[\alpha_k(X_1)\left|K_{12}\right| \mid Z_2\right]\right\} \\
&= E\left\{N(Y_2, X_2)\alpha_k(X_2) f(X_2)\right\}\int \left|K(u)\right| du + o(1) \\
&= O(1).
\end{aligned}$$

Hence, we obtain $V_{1n}^{(k)} = O_p\left(n^{-1}\right)$ by Assumption $A1$.

Using a mean value theorem argument,

$$V_{2n}^{(k)} = \left(\hat{\theta}_n - \theta_0\right)' S_{3n}\left(\hat{\theta}_n - \theta_0\right),$$

where

$$S_{3n} = \frac{1}{n(n-1)h^q}\sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n}\left[\dot{\psi}_i^{(k)} + \ddot{\psi}_i^{(k)}\left(\bar{\theta}_n\right)\left(\hat{\theta}_n - \theta_0\right)\right]\left[\dot{\psi}_j^{(k)} + \ddot{\psi}_j^{(k)}\left(\bar{\bar{\theta}}_n\right)\left(\hat{\theta}_n - \theta_0\right)\right]' K_{ij},$$

$$\left\|\bar{\theta}_n - \theta_0\right\| \leq \left\|\hat{\theta}_n - \theta_0\right\| \text{ and } \left\|\bar{\bar{\theta}}_n - \theta_0\right\| \leq \left\|\hat{\theta}_n - \theta_0\right\|$$

We can decompose S_{3n} in different terms, and reasoning as before it is easy to check that $S_{3n} = O_p(1)$, so that $V_{2n}^{(k)} = O_p\left(n^{-1}\right)$. ■

PROOF OF THEOREM 1

By Lemma 1,

$$nh^{q/2}T_n = nh^{q/2}\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} h^{-q}\sum_{k=1}^{m}\psi_1^{(k)}\psi_2^{(k)}K_{12} + O\left(h^{q/2}\right). \quad (9)$$

Therefore, by Assumption B, it suffices to establish the asymptotic normality of the first term in (9). Lemma 3 is useless now, because we have a degenerated U-statistic. Thus, we make use of a result for degenerate U-statistic, which is stated below.

Lemma 4 *(Hall, 1984) Let U_n be as in Lemma 3, where $E\left[H_n(Z_i, Z_j) \mid Z_i\right] = 0$ a.s. Define $G_n(Z_1, Z_2) = E\left[H_n(Z_3, Z_1)H_n(Z_3, Z_2) \mid Z_1, Z_2\right]$. If*

$$\lim_{n\to\infty}\frac{E\left[G_n(Z_1, Z_2)^2\right] + n^{-1}E\left[H_n(Z_1, Z_2)^4\right]}{\left\{E\left[H_n(Z_1, Z_2)^2\right]\right\}^2} = 0,$$

then

$$\frac{nU_n}{2\left\{E\left[H_n(Z_1, Z_2)^2\right]\right\}^{1/2}} \xrightarrow{d} N(0, 1).$$

In our case, $H_n(Z_1, Z_2) = h^{-q}\sum_{k=1}^{m}\psi_1^{(k)}\psi_2^{(k)}K_{12}$. First,

$$E\left[G_n(Z_1, Z_2)^2\right] = \sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{k'=1}^{m}\sum_{l'=1}^{m}\lambda_{kk'll'}$$

where each $\lambda_{kk'll'}$ is equal to

$$
\begin{aligned}
&h^{-4q}E\left\{\sigma_{kk'}(X_1)\sigma_{ll'}(X_2)E\left[K_{13}K_{23}\sigma_{kl}(X_3)\mid X_1,X_2\right]E\left[K_{13}K_{23}\sigma_{k'l'}(X_3)\mid X_1,X_2\right]\right\}\\
&=\frac{1}{h^{2q}}E\left\{\sigma_{kk'}(X_1)\sigma_{ll'}(X_2)\left[\int K(u)K\left(u+\frac{X_2-X_1}{h}\right)\sigma_{kl}(X_1+hu)f(X_1+hu)\,du\right]\right.\\
&\left.\left[\int K(u')K\left(u'+\frac{X_2-X_1}{h}\right)\sigma_{k'l'}(X_1+hu')f(X_1+hu')\,du'\right]\right\}\\
&=\frac{1}{h^{q}}E\left\{\sigma_{kk'}(X_1)\sigma_{ll'}(X_1+hv)\left[\int K(u)K(u+v)\sigma_{kl}(X_1+hu)f(X_1+hu)\,du\right]\right.\\
&\left.\left[\int K(u')K(u'+v)\sigma_{k'l'}(X_1+hu')f(X_1+hu')\,du'\right]f(X_1+hv)\,dv\right\}\\
&=\frac{1}{h^{q}}E\left\{\sigma_{kk'}(X_1)\sigma_{ll'}(X_1)\sigma_{kl}(X_1)\sigma_{k'l'}(X_1)f^3(X_1)\right\}\\
&\left[\int\int\int K(u)K(u+v)K(u')K(u'+v)\,du\,du'\,dv\right]+o\left(h^{-q}\right)\\
&=O\left(\frac{1}{h^{q}}\right),
\end{aligned}
$$

using Assumption S. Second,

$$
\begin{aligned}
&E\left[H_n(Z_1,Z_2)^2\right]\\
&=\sum_{k=1}^{m}\sum_{l=1}^{m}\frac{1}{h^{2q}}E\left(K\left(\frac{X_1-X_2}{h}\right)^2\sigma_{kl}(X_1)\sigma_{kl}(X_2)\right)\\
&=\sum_{k=1}^{m}\sum_{l=1}^{m}\frac{1}{h^{q}}E\left(\sigma_{kl}(X_1)\int K(u)^2\sigma_{kl}(X_1+hu)f(X_1+hu)\,du\right)\\
&=\sum_{k=1}^{m}\sum_{l=1}^{m}\frac{1}{h^{q}}E\left[\sigma_{kl}^2(X_1)f(X_1)\right]\int K(u)^2\,du+o\left(h^{-q}\right)=O\left(\frac{1}{h^{q}}\right).
\end{aligned}
$$

Third, denoting $K[(X_1-X_2)/h]$ by K_{12},

$$
\begin{aligned}
&E\left[H_n(Z_1,Z_2)^4\right]\\
&=\sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{k'=1}^{m}\sum_{l'=1}^{m}\frac{1}{h^{4q}}E\left[K_{12}^4\psi_1^{(k)}\psi_2^{(k)}\psi_1^{(k')}\psi_2^{(k')}\psi_1^{(l)}\psi_2^{(l)}\psi_1^{(l')}\psi_2^{(l')}\right]\\
&\leq\sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{k'=1}^{m}\sum_{l'=1}^{m}\frac{1}{h^{3q}}\prod_{p\in\{k,l,k',l'\}}\left\{E\left[h^{-q}K_{12}^4\sigma_p^4(X_1)\sigma_p^4(X_2)\right]\right\}^{1/4}\\
&\leq\sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{k'=1}^{m}\sum_{l'=1}^{m}\frac{1}{h^{3q}}\prod_{p\in\{k,l,k',l'\}}\left\{E\left[\sigma_p^4(X_1)\sigma_p^4(X_1)f(X_1)\right]\right\}^{1/4}+o\left(h^{-3q}\right)\\
&\leq O\left(\frac{1}{h^{3q}}\right).
\end{aligned}
$$

Therefore,

$$\lim_{n\to\infty} \frac{E\left[G_n(Z_1,Z_2)^2\right] + n^{-1}E\left[H_n(Z_1,Z_2)^4\right]}{\left\{E\left[H_n(Z_1,Z_2)^2\right]\right\}^2} = \lim_{n\to\infty}\left(h^q + (nh^q)^{-1}\right) = o(1),$$

by Assumption B, and Lemma 4 allows to conclude. ■

PROOF OF COROLLARY 1

Let us first consider the properties of T_n under H_1. Notice that $T_n = \sum_{k=1}^m T_n^{(k)}$, where

$$T_n^{(k)} = \binom{n}{2}^{-1} \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} H_n^{(k)}(Z_i,Z_j) + 2V_{1n}^{(k)} + V_{2n}^{(k)},$$

and $H_n^{(k)}(Z_i,Z_j) = h^{-q}\psi_i^{(k)}(\theta_1)\,\psi_j^{(k)}(\theta_1)\,K_{ij}$,

$$V_{1n}^{(k)} = \frac{1}{n(n-1)h^q}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \psi_i^{(k)}(\theta_1)\left(\hat{\psi}_j^{(k)} - \psi_j^{(k)}(\theta_1)\right)K_{ij},$$

$$V_{2n}^{(k)} = \frac{1}{n(n-1)h^q}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \left(\hat{\psi}_i^{(k)} - \psi_i^{(k)}(\theta_1)\right)\left(\hat{\psi}_j^{(k)} - \psi_j^{(k)}(\theta_1)\right)K_{ij}.$$

Reasoning as in the proof of Lemma 1 and using $A1$ with θ_1 in place of θ_0, we can easily show that $V_{1n}^{(k)} = o_p(1)$ and $V_{2n}^{(k)} = o_p(1)$, all $k = 1,\dots,m$. Now using Lemma 3 and a central limit theorem argument, we have

$$\begin{aligned}
&\sum_{k=1}^{m}\binom{n}{2}^{-1}\sum_{i=1}^{n}\sum_{j=i+1}^{n-1} H_n^{(k)}(Z_i,Z_j)\\
&= \sum_{k=1}^{m} E\left[H_n^{(k)}(Z_i,Z_j)\right] + O_p\left(n^{-1/2}\right)\\
&= \sum_{k=1}^{m} E\left[E^2\left[\psi^{(k)}(Y,X;\theta_1)\,|X\right] f(X)\right]\int K(u)\,du + O_p\left(n^{-1/2}\right).
\end{aligned}$$

This shows that T_n converges to a strictly positive limit under H_1.

By a similar reasoning, it is easily shown that, either under H_0 or H_1,

$$\hat{V}_n = \frac{2}{n(n-1)h^q}\sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \widehat{\psi}_i^{(k)}\widehat{\psi}_i^{(l)}\widehat{\psi}_j^{(k)}\widehat{\psi}_j^{(l)} K_{ij}^2 = V + o_p(1).$$

Corollary 1 then follows. ■

PROOF OF LEMMA 2

Lemma 2 follows straightforwardly after applying a mean value theorem argument, the Lindeberg-Levy central limit theorem and the strong law of large numbers. ■

PROOF OF THEOREM 2

The convergence of the finite dimensional distributions of $R_n(.)$ follows from Lemma 2 and applying the Linderberg-Levy central limit theorem. Thus, the theorem follows by proving tightness.

Let $D_1 = (s,t] = \times_{k=1}^{q}(s_j,t_j]$, $D_2 = (s',t'] = \times_{k=1}^{q}(s'_j,t'_j]$ two neighbors intervals in $\mathbb{R}^q$, i.e., they abut and for some $j \in \{1,2,..,q\}$, 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 process on $D[\mathbb{R}^q]$. Define

$$W_n(D_1) = \sum_{e_1=0,1} \cdots \sum_{e_q=0,1} (-1)^{q-\sum_p e_p} W_n\left(s_1 + e_1(t_1 - s_1), \cdots, s_q + e_q(t_q - s_q)\right). \tag{10}$$

By Condition (2.1.8) in Gaenssler and Stute (1979), a sufficient condition for tightness in $D[\mathbb{R}^q]$ is

$$\Pr\left(|W_n(D_1)| > \delta; |W_n(D_2)| > \delta\right) \leq C\delta^{-a}\left(\mu\left(D_1 \cup D_2\right)\right)^b \tag{11}$$

where $\mu(\cdot)$ is an arbitrary measure with continuous marginals and a, b and C are arbitrary constants such that $b > 1$ and $C \geq 0$. Using Markov inequality, a sufficient condition for (11) is

$$E\left(|W_n(D_1)|\,|W_n(D_2)|\right)^2 \leq C\left(\mu\left(D_1 \cup D_2\right)\right)^b. \tag{12}$$

Without loss of generality we will prove tightness for $q = 2$. From Lemma 2, we can write

$$R_n(t) = R_n^0(t) + R_n^1(t),$$

where

$$R_n^0(t) = \frac{1}{n}\sum_{i=1}^{n}\psi_i\Delta_i(t)$$

$$R_n^1(t) = E\left[\dot{\psi}_i(\theta_0)\,\Delta_i(t)\right]\frac{1}{n}\sum_{i=1}^{n} l\left(Y_i, X_i; \theta_0\right)$$

with $k-$th coordinate denoted by $R_n^{j(k)}, j = 0, 1$, i.e. $R_n^{0(k)}(t) = \frac{1}{n}\sum_{i=1}^n \psi_i^{(k)}\Delta_i(t)$, and $R_n^{1(k)}(t) = E\left[\dot{\psi}_i^{(k)}(\theta_0)\Delta_i(t)\right]\frac{1}{n}\sum_{i=1}^n l(Y_i, X_i; \theta_0)$, where $\dot{\psi}_i^{(k)}(\theta_0)$ is the k-th row in $\dot{\psi}_i(\theta_0)$. Then, applying definition in (10) to $R_n^{0(k)}(.)$,

$$\begin{aligned}\sqrt{n}R_n^{0(k)}(D_1) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_i^{(k)}(\Delta_i(t_1, t_2) - \Delta_i(s_1, t_2) - \Delta_i(t_1, s_2) + \Delta_i(s_1, s_2)) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \psi_i^{(k)}\Delta_i(D_1), \qquad (13)\end{aligned}$$

where $\Delta_i(D_j) = 1(X_i \in D_j)$. Lemma 5.1 in Stute (1997) 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

$$E\left(\left(\sum_{i=1}^n \alpha_i\right)^2\left(\sum_{i=1}^n \beta_i\right)^2\right) \le nE(\alpha_1^2\beta_1^2) + 3n(n-1)E(\alpha_1^2)E(\beta_1^2). \qquad (14)$$

Let $\alpha_i = \psi_i^{(k)}\Delta_i(D_1)$ and $\beta_i = \psi_i^{(k)}\Delta_i(D_2)$, then $\alpha_1^2\beta_1^2 = 0$ and applying (14) we get

$$\begin{aligned}n^2E\left(\left(R_n^{0(k)}(D_1)\right)^2\left(R_{ln}^{0(k)}(D_2)\right)^2\right) &\le \frac{3n(n-1)}{n^2}E\left(\psi_1^{(k)}\Delta_1(D_1)\right)^2E\left(\psi_1^{(k)}\Delta_1(D_2)\right)^2 \\ &\le 3\left(E\left(\left[\psi_1^{(k)}\right]^2\Delta_1(D_1 \cup D_2)\right)\right)^2.\end{aligned}$$

Then (12) holds for $R_n^{0(k)}(.)$ and this process is tight for arbitrary $k = 1, .., m$. Because the index parameter in $R_n^{1(k)}(t)$ is included in a deterministic continuous bounded function, it is straightforward to check that it is tight.

Recall that a sequence of stochastic process $\{W_n, n = 1, ..., +\infty\}$ is said to be tight if and only if for every $\varepsilon > 0$, there exists a compact set K such that $\sup_n \Pr\{W_n \in K\} > 1 - \varepsilon$, i.e. if there exists a compact set of the sample space where the process is included, with arbitrary high probability uniformly in n. Let $K^{j(k)}$ be the compact sets that includes each $R_n^{j(k)}(.)$ with arbitrary high probability, $j = 0, 1$, $k = 1, ..., m$. By Tychonoff Theorem (see Dudley, 1989, Th 2.2.8), the set $K^{(k)} = \{c^{(k)} = (c^0, c^1) : c^j \in K^{j(k)},\ j = 1, 2\}$ is compact within the product topology. Because summation is a continuous operator in $D[\mathbb{R}^q]$, it preserves compactness and the set $K^{(k)+} = \{c^{(k)+} = c^0 + c^1 : (c^0, c^1) \in K^{(k)}\}$ is compact. Therefore, the process $R_n^{(k)}(.)$ is tight. By Tychonoff Theorem again, the set $K = \{c = (c^{(1)+}, c^{(2)+}, .., c^{(m)+}) : c^{(k)+} \in K^{(k)+},\ k = 1, .., m\}$ is compact within the product topology. Since $R_n(t) \in K$ with arbitrary high probability uniformly in n, it is tight and the proof is completed. ■

PROOF OF THEOREM 3

Henceforth, $\psi_i^{*(k)} = V_i \psi_i^{(k)}(\theta_0)$ and $\hat{\psi}_i^{*(k)} = V_i \psi_i^{(k)}\left(\widehat{\theta}\right)$ and $E^*[.] \equiv E[.|\mathcal{Y}_n]$. Notice that $T_n^* = \sum_{k=1}^m T_n^{*(k)}$, where

$$T_n^{*(k)} = \widetilde{T}_n^{*(k)} + 2V_{1n}^{*(k)} + V_{2n}^{*(k)},$$

where

$$\widetilde{T}_n^{*(k)} = \binom{n}{2}^{-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} h_j^{-q} \psi_i^{*(k)} \psi_j^{*(k)} K_{ij},$$

$$V_{1n}^{*(k)} = \frac{1}{n(n-1)h^q} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \psi_i^{*(k)} \left(\widehat{\psi}_j^{*(k)} - \psi_j^{*(k)} \right) K_{ij},$$

$$V_{2n}^{*(k)} = \frac{1}{n(n-1)h^q} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left(\widehat{\psi}_i^{*(k)} - \psi_i^{*(k)} \right) \left(\widehat{\psi}_j^{*(k)} - \psi_j^{*(k)} \right) K_{ij}.$$

We will prove that $V_{1n}^{*(k)} = O_{p^*}\left(n^{-1}\right)$ and $V_{2n}^{*(k)} = O_{p^*}\left(n^{-1}\right)$ in probability for all $k = 1, ..., m$. Using a mean value theorem argument,

$$V_{1n}^{*(k)} = \left(\hat{\theta}_n - \theta_0\right)' S_{1n}^* + \left(\hat{\theta}_n - \theta_0\right)' S_{2n}^* \left(\hat{\theta}_n - \theta_0\right),$$

where

$$S_{1n}^* = \frac{1}{n(n-1)h^q} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \psi_i^{*(k)} \dot{\psi}_j^{*(k)} K_{ij},$$

$$S_{2n}^* = \frac{1}{n(n-1)h^q} \sum_{i=1}^{n} \sum_{\substack{j=1^i \\ j \neq i}}^{n} \psi_i^{*(k)} \ddot{\psi}_j^{*(k)} \left(\bar{\theta}_n\right) K_{ij}, \ \left\|\bar{\theta}_n - \theta_0\right\| \leq \left\|\hat{\theta}_n - \theta_0\right\|,$$

$\dot{\psi}_j^{*(k)} = \dot{\psi}_j^{(k)} V_j$, and $\ddot{\psi}_j^{*(k)}(\theta) = \ddot{\psi}^{(k)}\left(Y_j^*, X_j; \theta\right) V_j$. Since $E^*\left(V_j V_{j'}\right) = 0$ for $j \neq j'$, S_{1n}^* and S_{2n}^* are degenerate U-statistics. Hence, $E^*\left[S_{1n}^*\right] = 0$ and

$$E^*\left[S_{1n}^{*2}\right] = \left[\frac{1}{n(n-1)h^q}\right]^2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} E^*\left[\psi_i^{*(k)} \dot{\psi}_j^{*(k)} K_{ij}\right]^2$$

$$= \left[\frac{1}{n(n-1)h^q}\right]^2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left(\psi_i^{(k)}\right)^2 \left(\dot{\psi}_j^{(k)}\right)^2 K_{ij}.$$

Now, using similar arguments than in the proof of Lemma 1,

$$E\left[E^*\left[S_{1n}^{*2}\right]\right]=\left[\frac{1}{n\left(n-1\right)h^q}\right]^2 n(n-1)O(h^q)=O(n^{-2}h^{-q}),$$

so that $S_{1n}^*=O_{p^*}(n^{-1}h^{-q/2})$ in probability. Similarly, we can prove that $E^*\left|S_{2n}^*\right|=O_{p^*}(1)$ in probability, so that $S_{2n}^*=O_{p^*}(1)$ in probability. As $\left(\hat{\theta}-\theta_0\right)=O_p\left(n^{-1/2}\right)$, we get that $nh^{q/2}V_{1n}^{*(k)}=o_{p^*}(1)$ in probability. Similarly, $nh^{q/2}V_{2n}^{*(k)}=o_{p^*}(1)$ in probability.

Concerning asymptotic normality, we will treat the case where $m=1$ for the sake of simplicity, so that we consider

$$\widetilde{T}_n^*=\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}H_n^*\left(Z_i,Z_j\right)=\binom{n}{2}^{-1}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}h_j^{-q}\psi_i^*\psi_j^*K_{ij},$$

Notice that $E^*\left[H_n^*\left(Z_i,Z_j\right)|V_i\right]=0$, for all i. Let us define

$$\begin{aligned}\sigma_n^2 &\equiv E^*\left[\widetilde{T}_n^2\right]\\ &=\binom{n}{2}^{-2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\sum_{i'=1}^{n-1}\sum_{j'=i'+1}^{n}E^*\left[H_n^*\left(Z_i,Z_j\right)H_n^*\left(Z_{i'},Z_{j'}\right)\right]\\ &=\binom{n}{2}^{-2}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}h_j^{-2q}\left(\psi_i\right)^2\left(\psi_j\right)^2K_{ij}^2=\frac{1}{n(n-1)h^q}\widehat{V}_n,\end{aligned}$$

$$\begin{aligned}G_1 &\equiv \binom{n}{2}^{-4}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}E^*\left[H_n^{*4}\left(Z_i,Z_j\right)\right]\\ &=E^2\left(V_1^4\right)\binom{n}{2}^{-4}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}h_j^{-4q}\psi_i^4\psi_j^4K_{ij}^4,\end{aligned}$$

$$\begin{aligned}G_2 &\equiv \binom{n}{2}^{-4}\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n}E^*\left[H_n^{*2}\left(Z_i,Z_j\right)H_n^{*2}\left(Z_i,Z_k\right)\right]\\ &=E\left(V_1^4\right)\binom{n}{2}^{-4}\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n}h_j^{-4q}\psi_i^4\psi_j^2\psi_k^2K_{ij}^2K_{ik}^2,\end{aligned}$$

$$\begin{aligned}G_3 &\equiv \binom{n}{2}^{-4}\sum_{i=1}^{n-3}\sum_{j=i+1}^{n-2}\sum_{k=j+1}^{n-1}\sum_{l=k+1}^{n}E^*\left[H_n^*\left(Z_i,Z_j\right)H_n^*\left(Z_i,Z_k\right)H_n^*\left(Z_l,Z_j\right)H_n^*\left(Z_l,Z_k\right)\right]\\ &=\binom{n}{2}^{-4}\sum_{i=1}^{n-3}\sum_{j=i+1}^{n-2}\sum_{k=j+1}^{n-1}\sum_{l=k+1}^{n}h_j^{-4q}\psi_i^2\psi_j^2\psi_k^2\psi_l^2K_{ij}K_{ik}K_{lj}K_{lk}.\end{aligned}$$

By Proposition 3.2 in De Jong (1987), we have $\sigma_n^{-1}\widetilde{T}_n^* \xrightarrow{d^*} N(0,1)$ in probability if G_1, G_2 and G_3 are of lower order in probability than $\left(\sigma_n^2\right)^2$. Now, G_1, G_2 and $\left(\sigma_n^2\right)^2$ are positive and it is not difficult to check, as in the proof of Theorem 1, that

$$\begin{aligned}
E\left[\left(\sigma_n^2\right)^2\right] &= E\left[\binom{n}{2}^{-2}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} h_j^{-2q}\left(\psi_i\right)^2\left(\psi_j\right)^2 K_{ij}^2\right]^2 \\
&= n^{-8}O\left[n^2h^{-3q}+n^3h^{-2q}+n^4h^{-2q}\right] = O(n^{-4}h^{-2q}), \\
E\left[G_1\right] &= O\left(n^{-6}h^{-3q}\right), \\
E\left[G_2\right] &= O\left(n^{-5}h^{-2q}\right), \\
E\left|G_3\right| &= O\left(n^{-4}h^{-q}\right).
\end{aligned}$$

Hence, the condition is fulfilled and we get the desired result. It is straightforward while cumbersome to check the result for arbitrary m. ■

PROOF OF THEOREM 4

Using the independence of the sequence $\{V_i\}$, we obtain the covariance structure

$$nE^*\left(R_n^*(x_1)\, R_n^*(x_2)'\right) = \frac{1}{n}\sum_{i=1}^{n}\hat{r}_i(x_1)\,\hat{r}_i(x_2)',$$

where $\hat{r}_i(x) = \hat{\psi}_i\Delta_i(x) + \left(\frac{1}{n}\sum_{i=1}^{n}\dot{\psi}_i\left(\hat{\theta}_n\right)\Delta_i(x)\right) l\left(Y_i, X_i; \hat{\theta}_n\right)$. By A2, Slutsky theorem and the strong law of large numbers,

$$nE^*\left(R_n^*(x_1)\, R_n^*(x_2)'\right) \xrightarrow{a.s.} E\left(r_i(x_1)\, r_i(x_2)'\right),$$

so the process has the same covariance structure. Let $\lambda \in \mathbb{R}^m$ such that $\|\lambda\| = 1$. To obtain gaussianity, we check the conditions of the Lindeberg-Feller central limit theorem. Define

$$\hat{r}_i^*(x) = \sigma(x)\, b_n(x)\, d_{in}^*(x), \text{ where } d_{in}^*(x) = n^{-1/2}\lambda'\hat{r}_i(x)\, V_i/\hat{\sigma}_n(x),$$

$b_n(x) = \hat{\sigma}_n(x)/\sigma(x)$, $\hat{\sigma}_n^2(x) = \lambda' n^{-1}\sum_{i=1}^{n}\hat{r}_i(x)\hat{r}_i(x)'\lambda$ and $\sigma^2(x) = \lambda' E\left(r_i(x)\, r_i(x)'\right)\lambda$. The triangular array $\{d_{in}^*(x) : i = 1, ..., n,\ n = 1, ... + \infty\}$ has rows elements that are independent conditionally on $\mathcal{Y}_n$, with

$$E^*\left[d_{in}^*(x)\right] = n^{-1/2}\lambda'\hat{r}_i(x)\, E^*\left[V_i\right]/\hat{\sigma}_n(x) = 0 \text{ a.s.}$$

and

$$Var^*\left[\sum_{i=1}^{n} d_{in}^*(x)\right] = \hat{\sigma}_n^{-2}(x)\,\lambda' n^{-1}\sum_{i=1}^{n}\hat{r}_i(x)\,\hat{r}_i'(x)\,\lambda Var^*[V_i] = 1 \text{ a.s.}$$

Let $k_V = \sup|V_i|$. We can check that the Lindeberg condition holds for $d_{in}^*(x)$ with probability one as follows

$$\begin{aligned}
&\sum_{i=1}^{n} E^*\left[d_{in}^{*2}(x)\,1\left(|d_{in}^*(x)|>\delta\right)\right] \\
&= \frac{1}{n}\hat{\sigma}_n^{-2}(x)\,\lambda'\sum_{i=1}^{n}\hat{r}_i(x)\,\hat{r}_i'(x)\,\lambda E^*\left[V_i 1\left(\left|n^{-1/2}\lambda'\hat{r}_i(x)\,V_i/\hat{\sigma}_n(x)\right|>\delta\right)\right] \\
&\leq \frac{1}{n}\hat{\sigma}_n^{-2}(x)\,\lambda'\sum_{i=1}^{n}\hat{r}_i(x)\,\hat{r}_i'(x)\,\lambda k_V E^*\left[1\left(\left|n^{-1/2}\lambda'\hat{r}_i(x)/\hat{\sigma}_n(x)\right|>\delta\right)\right] \\
&\leq \frac{1}{n}\hat{\sigma}_n^{-2}(x)\,\lambda'\sum_{i=1}^{n}\hat{r}_i(x)\,\hat{r}_i'(x)\,\lambda k_V 1\left(\left|\lambda'\hat{r}_i(x)/\hat{\sigma}_n(x)\right|>\delta n^{1/2}\right) \\
&= \frac{1}{n}\hat{\sigma}_n^{-2}(x)\,\lambda'\sum_{i=1}^{n}r_i(x)\,r_i'(x)\,\lambda k_V 1\left(\left|\lambda' r_i(x)/\sigma_n(x)\right|>\delta n^{1/2}\right)+o_p(1)\ a.s. \\
&= o_p(1)\ a.s.
\end{aligned}$$

Hence $\sum_{i=1}^{n} d_{in}^*(x) \xrightarrow{d} N(0,I_m)$ almost surely. Since $b_n(x) \xrightarrow{a.s.} 1$, applying Slutsky theorem gives that $R_n^*(x_1) \xrightarrow{d^*} N(0,\sigma(x))$ with probability one. The convergence of the vector $(R_n^*(x_1),..,R_n^*(x_s))$ can be proved using analogous methods. Therefore, the convergence of finite dimensional distributions follows almost surely.

Tightness of the $k-$th coordinate of $R_n(x)$ is shown using Condition (2.1.8) in Gaenssler and Stute (1979). We define the vector

$$R_n^{0*}(x) = \frac{1}{n}\sum_{i=1}^{n}\hat{\psi}_i\Delta_i(x)\,V_i$$

with $k-$th coordinate denoted by $R_n^{0(k)*}(x)$. Applying Equation (10) as in (13) we denote

$$\sqrt{n}R_n^{0(k)*}(D_j) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\psi}_i^{(k)}\Delta_i(D_j)\,V_i, \quad j=1,2.$$

where D_1, D_2 are two neighbors intervals. Thus a slight extension of the arguments in Lemma 5.1 in Stute (1997) yields

$$
\begin{aligned}
n^2 E^*\left(R_n^{0(k)*}(D_1)\, R_n^{0(k)*}(D_2)\right)^2 &\le \frac{C}{n^2}\sum_{\substack{i,j=1\\ i\neq j}}^{n} E^*\left(\hat{\psi}_i^{(k)}\Delta_i(D_1)\, V_i\right)^2 E^*\left(\hat{\psi}_j^{(k)}\Delta_i(D_2)\, V_j\right)^2 \\
&\le \left[\frac{C}{n}\sum_{i=1}^{n} E^*\left(\hat{\psi}_i^{(k)}\Delta_i(D_1\cup D_2)\, V_i\right)^2\right]^2 \\
&= \left[\frac{C}{n}\sum_{i=1}^{n} \hat{\psi}_i^{(k)2}\Delta_i(D_1\cup D_2)\right]^2 \\
&\overset{a.s.}{\to} C\left[E\left(\psi_i^{(k)2}\Delta_i(D_1\cup D_2)\right)\right]^2,
\end{aligned}
$$

for some constant $C > 0$ and (12) holds for $R_n^{0(k)*}(.)$ with probability one. Using the same arguments as in Theorem 2, tightness of the $k-$th coordinate is extended to the whole process $R_n^*(x)$. ■

References

[1] Andrews, D. W. K. (1997): "A Conditional Kolmogorov Test," Econometrica, Vol. 65, pp. 1097-1128.

[2] Bierens, H. (1982): "Consistent Model Specification Tests," Journal of Econometrics, Vol. 20, pp. 105-134.

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

[4] Bierens, H. and W. Ploberger (1997): "Asymptotic Theory of Integrated Conditional Moment Tests," Econometrica, Vol. 65, pp. 1129-1151.

[5] Burbidge, J. B., L. Magee and A. L. Robb (1988): "Alternative Transformations to Handle Extreme Values of the Dependent Variable," Journal of the American Statistical Association, Vol. 83, pp. 123-127.

[6] De Jong, P. (1987): "A Central Limit Theorem for Generalized Quadratic Forms," Probability Theory and Related Fields, Vol. 75, pp. 261–277.

[7] De Jong, R.M. (1996): "On the Bierens Test Under Data Dependence," Journal of Econometrics, Vol. 72, pp. 1-32.

[8] 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, Vol. 9, pp. 70-90.

[9] Delgado, M (1993): "Testing the Equality of Nonparametric Regression Curves," Statistics and Probability Letters, Vol. 17, pp. 199-204.

[10] Delgado, M and M. Domínguez (1995): "A Consistent Test of Significance," Working Paper Universidad Carlos III de Madrid.

[11] Delgado, M and M. Domínguez (1997): "Consistent Specification Testing of Quantile Regression Models," Working Paper Universidad Carlos III de Madrid.

[12] Diebolt, J. (1995): "A Nonparametric Test for the Regression Function: Asymptotic Theory," Journal of Statistical Planning and Inference, Vol. 44, pp. 1-17.

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

[14] Fan, Y. (1994); "Testing the Goodness-of-Fit of a Parametric Density Function By Kernel Method," Econometric Theory, Vol. 10, pp. 316–356.

[15] Fan, Y. and Q. Li (1996): "Consistent Model Specification Tests: Nonparametric versus Bierens' Tests," Unpublished manuscript.

[16] Fan, Y. and Q. Li (1996): "Consistent Model Specification Tests: Omitted Variables, Parametric and Semiparametric Functional Forms," Econometrica, 64(4), pp. 865–890..

[17] Gaenssler, P. and W. Stute (1979): "Empirical Processes: A Survey of Results for Independent and Identically Distributed Random Variables," Annals of Probability, Vol. 7, pp. 193-243.

[18] Gozalo, P. (1993): "A Consistent Model Specification Test for Nonparametric Estimation of Regression Function Models," Econometric Theory, Vol. 9, pp. 451-477

[19] Hall, P. (1984): "Central Limit Theorem for Integrated Square Error of Multivariate Nonparametric Density Estimators," Journal of Multivariate Analysis, Vol. 14, pp. 1-16.

[20] Hall, P. and J. L. Horowitz (1996): " Bootstrap Critical Values for Tests Based on Generalized-Method-of Moments Estimators," Econometrica, Vol. 64, pp. 891-916.

[21] Hansen, B. (1996): "Inference When a Nuisance Parameter is Not Identified Under the Null Hypothesis," Econometrica, Vol. 64, pp. 413-430.

[22] Härdle, W. and E. Mammen (1993): "Comparing Nonparametric versus Parametric Regression Fits," Annals of Statistics, Vol. 21, pp. 1926-1947.

[23] Hong, Y. and H. White (1995): "Consistent Specification Testing via Nonparametric Series Regressions," Econometrica, Vol. 63, pp. 1133-1160.

[24] Hong-Zhy, A and C. Bing (1991): "A Kolmogorov-Smirnov Type Statistic with Application to Test for Nonlinearity in Time Series," International Statistical Review, Vol. 59, pp. 287-307.

[25] Horowitz J., and W. Härdle (1993): "Testing a Parametric Model Against a Semiparametric Alternative," Econometric Theory, Vol. 10, pp.821-848.

[26] Lewbel, A. (1995): "Consistent Nonparametric Hypothesis Tests with an Application to Slutsky Symmetry," Journal of Econometrics, Vol. 67, pp. 379-401.

[27] Li Q. and S. Wang (1996): " A Simple Consistent Bootstrap Test for a Parametric Regression Function," Unpublished Manuscript.

[28] Newey, W. K. (1985): "Maximum Likelihood Specification Testing and Conditional Moment Tests," Econometrica, Vol. 53, pp. 1047-1070.

[29] Powell, J. L., J. H. Stock and T. M. Stocker (1989): "Semiparametric Estimation of Index Coefficients," Econometrica, Vol. 57, pp. 1403-1430.

[30] Robinson, P. (1991). "Best Nonlinear Three-Stage Least Squares Estimation of Certain Econometric Models," Econometrica, Vol. 59. pp. 755-786.

[31] Rodriguez-Campos, M.C., W. Gonzalez-Manteiga and R. Cao (1997): "Testing the Hypothesis of a Generalized Linear Regression Model Using Nonparametric Regression Estimation," Journal of Statistical Planning and Inference, forthcoming.

[32] Rosenblatt, M. (1975): "A Quadratic Measure of Deviations of Two-dimensional Density Estimates an a Test of Independence," Annals of Statistics, Vol. 3, pp. 1-14.

[33] Stute, W. (1997): "Nonparametric Model Checks for Regression," Annals of Statistics, Vol. 25, pp. 613-641.

[34] Stute, W., W. Gonzalez-Manteiga and M. Presedo (1996): "Bootstrap Approximations in Model Checks for Regression," Journal of the American Statistical Association, Vol. 93, pp. 141-149.

[35] 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, Vol. 86, pp. 420-426.

[36] Tauchen, G. (1985): "Diagnostic Testing and Evaluation of Maximum Likelihood Models," Journal of Econometrics, Vol. 30, pp. 415-443

[37] Wu, C. F. J. (1986): "Jacknife, Bootstrap and other Resampling Methods in Regression Analysis (with discussion)," Annals of Statistics, Vol. 14 pp. 1261-1350.

[38] Zheng, X. (1996): "A Consistent Test of Functional Form via Nonparametric Estimation Technique," Journal of Econometrics, Vol. 75, pp. 263-289.

Table 1:
Empirical frequencies of rejections under DGP_0.

d		n=50	n=100	n=250		n=50	n=100	n=250		n=50	n=100	n=250
2.0	t_n	0.027 0.022	0.014 0.007	0.026 0.012	t_n^*	0.027 0.022	0.013 0.005	0.021 0.008	t_n^{**}	0.099 0.059	0.112 0.055	0.102 0.058
1.5	t_n	0.037 0.028	0.025 0.015	0.032 0.020	t_n^*	0.035 0.026	0.024 0.011	0.034 0.016	t_n^{**}	0.099 0.058	0.106 0.051	0.110 0.051
1.0	t_n	0.049 0.034	0.041 0.025	0.049 0.029	t_n^*	0.047 0.031	0.041 0.015	0.047 0.019	t_n^{**}	0.102 0.055	0.104 0.049	0.112 0.059
0.5	t_n	0.073 0.044	0.070 0.040	0.068 0.039	t_n^*	0.070 0.038	0.065 0.032	0.068 0.031	t_n^{**}	0.104 0.050	0.100 0.057	0.115 0.059
0.1	t_n	0.102 0.055	0.092 0.052	0.080 0.041	t_n^*	0.089 0.051	0.085 0.045	0.077 0.035	t_n^{**}	0.096 0.050	0.098 0.049	0.102 0.046
0.05	t_n	0.106 0.046	0.104 0.052	0.091 0.044	t_n^*	0.093 0.050	0.096 0.048	0.087 0.037	t_n^{**}	0.097 0.049	0.103 0.052	0.104 0.050
0.025	t_n	0.119 0.040	0.101 0.049	0.108 0.056	t_n^*	0.109 0.066	0.094 0.048	0.103 0.051	t_n^{**}	0.104 0.055	0.100 0.052	0.118 0.058
					c_n^*	0.065 0.037	0.073 0.032	0.095 0.046	c_n^{**}	0.095 0.056	0.104 0.047	0.104 0.053

Table 2

Empirical frequencies of rejections under DGP_1 ($\delta = 1$).

d		n=50	n=100	n=250		n=50	n=100	n=250		n=50	n=100	n=250
2.0	t_n	0.902 0.858	1.000 1.000	1.000 1.000	t_n^*	0.898 0.814	1.000 1.000	1.000 1.000	t_n^{**}	0.943 0.890	0.999 0.996	1.000 1.000
1.5	t_n	0.959 0.926	1.000 1.000	1.000 1.000	t_n^*	0.952 0.899	1.000 1.000	1.000 1.000	t_n^{**}	0.978 0.945	1.000 0.999	1.000 1.000
1.0	t_n	0.970 0.950	1.000 1.000	1.000 1.000	t_n^*	0.967 0.922	1.000 1.000	1.000 1.000	t_n^{**}	0.981 0.955	1.000 1.000	1.000 1.000
0.5	t_n	0.949 0.912	0.998 0.998	1.000 1.000	t_n^*	0.941 0.897	0.998 0.997	1.000 1.000	t_n^{**}	0.953 0.912	0.999 0.998	1.000 1.000
0.1	t_n	0.843 0.783	0.986 0.973	1.000 1.000	t_n^*	0.826 0.764	0.984 0.966	1.000 1.000	t_n^{**}	0.820 0.737	0.983 0.969	1.000 1.000
0.05	t_n	0.734 0.602	0.954 0.935	1.000 0.998	t_n^*	0.720 0.625	0.952 0.928	1.000 0.998	t_n^{**}	0.700 0.572	0.955 0.931	1.000 0.998
0.025	t_n	0.594 0.361	0.920 0.871	0.996 0.993	t_n^*	0.596 0.486	0.913 0.874	0.996 0.992	t_n^{**}	0.536 0.373	0.908 0.860	0.995 0.992
					c_n^*	0.139 0.105	0.234 0.151	0.801 0.518	c_n^{**}	0.846 0.753	0.924 0.869	0.987 0.968

Table 3:
Empirical frequencies of rejections under DGP_2 ($\delta = 2$).

d		n=50	n=100	n=250		n=50	n=100	n=250		n=50	n=100	n=250
2.0	t_n	0.110 0.104	0.052 0.049	1.000 0.995	t_n^*	0.111 0.094	0.052 0.049	1.000 0.991	t_n^{**}	0.134 0.122	0.054 0.053	0.906 0.838
1.5	t_n	0.154 0.136	0.982 0.955	1.000 1.000	t_n^*	0.153 0.123	0.979 0.908	1.000 1.000	t_n^{**}	0.252 0.193	0.745 0.622	0.983 0.969
1.0	t_n	0.988 0.963	1.000 1.000	1.000 1.000	t_n^*	0.984 0.933	1.000 1.000	1.000 1.000	t_n^{**}	0.834 0.761	0.972 0.945	1.000 1.000
0.5	t_n	0.999 0.997	1.000 1.000	1.000 1.000	t_n^*	0.999 0.996	1.000 1.000	1.000 1.000	t_n^{**}	0.967 0.935	0.999 0.994	1.000 1.000
0.1	t_n	0.959 0.933	1.000 1.000	1.000 1.000	t_n^*	0.952 0.928	1.000 1.000	1.000 1.000	t_n^{**}	0.910 0.847	0.997 0.992	1.000 1.000
0.05	t_n	0.882 0.795	0.999 0.999	1.000 1.000	t_n^*	0.867 0.803	0.999 0.999	1.000 1.000	t_n^{**}	0.815 0.697	0.995 0.981	1.000 1.000
0.025	t_n	0.734 0.475	0.990 0.973	1.000 1.000	t_n^*	0.738 0.610	0.989 0.974	1.000 1.000	t_n^{**}	0.632 0.473	0.972 0.936	1.000 1.000
					c_n^*	0.157 0.132	0.417 0.159	0.986 0.948	c_n^{**}	0.459 0.298	0.678 0.552	0.907 0.838

Table 4:
Empirical frequencies of rejections under DGP_3 ($\delta = 3$).

d		n=50	n=100	n=250		n=50	n=100	n=250		n=50	n=100	n=250
2.0	t_n	0.124	0.064	0.006	t_n^*	0.124	0.064	0.006	t_n^{**}	0.229	0.090	0.006
		0.105	0.061	0.006		0.090	0.055	0.006		0.166	0.073	0.006
1.5	t_n	0.102	0.054	0.017	t_n^*	0.098	0.055	0.010	t_n^{**}	0.124	0.059	0.020
		0.091	0.052	0.011		0.071	0.047	0.010		0.103	0.053	0.018
1.0	t_n	0.151	0.914	1.000	t_n^*	0.144	0.909	1.000	t_n^{**}	0.214	0.711	0.979
		0.124	0.826	1.000		0.102	0.715	1.000		0.143	0.559	0.954
0.5	t_n	0.994	1.000	1.000	t_n^*	0.991	1.000	1.000	t_n^{**}	0.946	0.992	1.000
		0.975	1.000	1.000		0.963	1.000	1.000		0.884	0.984	1.000
0.1	t_n	0.968	1.000	1.000	t_n^*	0.964	1.000	1.000	t_n^{**}	0.943	0.998	1.000
		0.923	0.999	1.000		0.919	0.999	1.000		0.876	0.993	1.000
0.05	t_n	0.900	0.998	1.000	t_n^*	0.884	0.998	1.000	t_n^{**}	0.828	0.990	1.000
		0.792	0.996	1.000		0.799	0.995	1.000		0.722	0.982	1.000
0.025	t_n	0.723	0.983	1.000	t_n^*	0.728	0.979	1.000	t_n^{**}	0.643	0.961	0.961
		0.487	0.964	0.999		0.594	0.961	0.999		0.490	0.926	0.926
					c_n^*	0.132	0.135	0.901	c_n^{**}	0.265	0.435	0.890
						0.125	0.065	0.655		0.188	0.266	0.803