

Diack, Cheikh A. T.; Thomas-Agnan, Christine

Working Paper

A nonparametric test of the non-convexity of regression

SFB 373 Discussion Paper, No. 1998,43

Provided in Cooperation with:

Collaborative Research Center 373: Quantification and Simulation of Economic Processes,
Humboldt University Berlin

Suggested Citation: Diack, Cheikh A. T.; Thomas-Agnan, Christine (1998) : A nonparametric test of the non-convexity of regression, SFB 373 Discussion Paper, No. 1998,43, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, Berlin, <https://nbn-resolving.de/urn:nbn:de:kobv:11-10056968>

This Version is available at:

<https://hdl.handle.net/10419/61308>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

A NONPARAMETRIC TEST OF THE NON-CONVEXITY OF REGRESSION ¹

by

*Cheikh A.T. DIACK** and *Christine THOMAS-AGNAN ***

*Université Paul Sabatier Laboratoire de Statistique et Probabilités
118 route de Narbonne 31062 Toulouse, France.

**GREMAQ, Manufacture des tabacs, 21 Allées de Brienne,
31000 Toulouse, France.

Abstract

This paper proposes a nonparametric test of the non-convexity of a smooth regression function based on least squares or hybrid splines. By a simple formulation of the convexity hypothesis in the class of all polynomial cubic splines, we build a test which has an asymptotic size equal to the nominal level. It is shown that the test is consistent and is robust to nonnormality. The behavior of the test under the local alternatives is studied.

¹*key words: least squares estimator, test of convexity, B-splines, modulus of continuity*

1 INTRODUCTION

This paper proposes a test of non-convexity of an unknown regression function in a nonparametric regression model. Interest in this problem which has been addressed for example in Schlee(1980) and Yatchew(1992), arises in particular from econometric models. Economic theory predicts the convexity of functions like for example cost functions, production functions, Engel curves... Such tests provide a way of confronting theory with real data sets.

We consider the following regression model where we are given n realizations y_{ij} of real random variables

$$y_{ij} = f(x_i) + \varepsilon_{ij}, i = 1, \dots, r, j = 1, \dots, n_i \quad \text{with} \quad x_i \in (0, 1), i = 1, \dots, r.$$

At each deterministic design point x_i , ($i = 1, \dots, r$), n_i measurements are taken. The probability measure assigning mass $\mu_i = n_i/n$ to the point x_i ($\sum \mu_i = 1$) is referred to as the design and will be denoted by μ^n . We assume that the random errors ε_{ij} are uncorrelated and identically distributed with mean zero. Their variance σ^2 will be assumed known. Finally f is an unknown smooth regression function.

We use a cubic spline function to approximate the function f , not only because of the common use of cubic splines in approximation problems, but also because they allow to formulate these convexity hypothesis in a very simple way. Indeed, if g is a cubic spline then, its second derivative is a linear function between any pair of adjacent knots η_i and η_{i+1} , and it follows that g is a convex function in the interval $\eta_i \leq x \leq \eta_{i+1}$ if and only if $g''(\eta_i)$ and $g''(\eta_{i+1})$ are both non negative. This property was used by Dierckx (1980) to define a convex estimator. After some preliminaries, we rapidly sketch in the next section, the cubic spline estimator. In section 3, we describe our test of non-convexity of the regression function and the rest of the section is devoted to a discussion and demonstration of its properties. Section 4 presents an application to a real data set.

2 NOTATIONS AND PRELIMINARIES

2.1 The Cubic Spline Estimator

Let p be a positive continuous density on $(0,1)$. We assume that

$$\min_{0 \leq x \leq 1} p(x) > 0.$$

Let $\eta_0 = 0 < \eta_1 < \dots < \eta_{k+1} = 1$ be a subdivision of the interval $(0,1)$ by k distinct points defined by

$$\int_0^{\eta_i} p(x) dx = i/(k+1), \quad i = 0, \dots, k+1. \quad (2.1)$$

Let $\delta_k = \max_{0 \leq i \leq k} (\eta_{i+1} - \eta_i)$, we see that

$$\delta_k / \min_i (\eta_{i+1} - \eta_i) \leq \max_x p(x) / \min_x p(x). \quad (2.2)$$

For each fixed set of knots of the form (2.1), we define $\mathcal{S}(k,d)$ as the collection of all polynomial splines of order d (degree $\leq d-1$) having for knots $\eta_1 < \dots < \eta_k$. The class $\mathcal{S}(k,d)$ of such splines is a linear space of functions of dimension $(k+d)$. A basis for this linear space is provided by the B-splines (see Schumaker (1981)). Let $\{N_1, \dots, N_{k+d}\}$ the set of normalized B-splines associated to the following non-decreasing sequence $\{t_1, \dots, t_{k+2d}\}$:

$$\begin{cases} t_1 \leq t_2 \leq \dots \leq t_d = 0 \\ t_{2d+k} \geq t_{2d+k-1} \geq \dots \geq t_{d+k+1} = 1 \\ t_{d+l} = \eta_l \quad \text{for } l = 1, \dots, k \end{cases}$$

The reader is referred to Schumaker (1981) for a discussion of these B-splines.

In what follows, we shall only work with the class of cubic splines: $\mathcal{S}(k,4)$. A cubic spline on the interval $(0,1)$ is said to be a natural cubic spline if its second and third derivatives are zero at 0 and 1. They imply that a natural cubic spline is linear on the two extreme intervals $(0, \eta_1)$ and $(\eta_k, 1)$. It will be convenient to introduce the following notations:

$$N(x) = (N_1(x), \dots, N_{k+4}(x))' \in \mathbb{R}^{k+4}.$$

$$F = (N(x_1), \dots, N(x_r)); (k+4) \times r \quad \text{matrix.}$$

We will denote by \hat{f}_n the least squares spline estimator of f :

$$\hat{f}_n(x) = \sum_{p=1}^{k+4} \hat{\theta}_p N_p(x) \quad (2.3)$$

where

$$\hat{\Theta} = (\hat{\theta}_1, \dots, \hat{\theta}_{k+4})' = \arg \min_{\Theta \in \mathbb{R}^{k+4}} \sum_{i=1}^r \sum_{j=1}^{n_i} (y_{ij} - \sum_{p=1}^{k+4} \theta_p N_p(x_i))^2.$$

Let

$$\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}, \bar{\varepsilon}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \varepsilon_{ij}, \quad (2.4)$$

$$\bar{Y} = (\bar{y}_1, \dots, \bar{y}_r)', \bar{\varepsilon} = (\bar{\varepsilon}_1, \dots, \bar{\varepsilon}_r)', f_{\Delta} = (f(x_1), \dots, f(x_r))'$$

Let $\mathcal{D}(\mu^n)$ be the $r \times r$ diagonal matrix with diagonal elements μ_1, \dots, μ_r , then, basic least squares arguments prove that:

$$\hat{\Theta} = M^{-1}(\mu^n) F \mathcal{D}(\mu^n) \bar{Y} \quad \text{with} \quad M(\mu^n) = \sum_{i=1}^r N(x_i) N'(x_i) \mu_i = F \mathcal{D}(\mu^n) F'.$$

Asymptotic properties of this estimator have been established in Argarwal and Studden(1980).

We will introduce the hybrid cubic spline estimator of f denoted by \tilde{f}_n as in Kelly and Rice (1990). It is defined to be the cubic spline minimizing the penalized least squares criterion i.e:

$$\tilde{f}_n = \sum_{p=1}^{k+4} \tilde{\theta}_p N_p(x) \quad (2.5)$$

where

$$\tilde{\Theta} = \arg \min_{\Theta \in \mathbb{R}^{k+4}} \sum_{i=1}^r \sum_{j=1}^{n_i} (y_{ij} - \sum_{p=1}^{k+4} \theta_p N_p(x_i))^2 + \lambda \int_0^1 (\sum_{p=1}^{k+4} \theta_p N_p''(x))^2 dx \quad (2.6)$$

with a smoothing parameter $\lambda > 0$.

This estimator is therefore a hybrid of smoothing splines and least squares splines in the sense that the amount of smoothing is controlled by both a penalty term through the choice of λ as in smoothing splines and by the number of knots, k , as in least squares splines.

As above with $\hat{\Theta}$, we can specify $\tilde{\Theta}$. For this we need to introduce the following matrix W defined by

$$W_{ij} = \int_0^1 N_i''(x) N_j''(x) dx, \quad i, j = 1, \dots, k+4.$$

W is a band matrix. Indeed, $N_i(x) = 0$ if $x \notin (t_i, t_{i+4})$ (see Argarwal and Studden(1980)). Hence $W_{ij} = 0$ if $|i - j| \geq 4$.

It is easy to see that if $g = \sum_{p=1}^{k+4} \theta_p N_p(x)$ then

$$\int_0^1 (g''(x))^2 dx = \Theta' W \Theta. \quad (2.7)$$

Now, we are ready for the main result of this section.

Lemma 1 *If $\tilde{\Theta}$ satisfies (2.6) then*

$$\tilde{\Theta} = \Gamma^{-1}(\mu^n) F \mathcal{D}(\mu^n) \bar{Y} \quad \text{with} \quad \Gamma(\mu^n) = M(\mu^n) + \lambda W.$$

Proof

Let $g = \sum_{p=1}^{k+4} \theta_p N_p(x)$ be a cubic spline.

Let

$$L(g) = \sum_{i=1}^r \sum_{j=1}^{n_i} (y_{ij} - \sum_{p=1}^{k+4} \theta_p N_p(x_i))^2 + \lambda \int_0^1 (\sum_{p=1}^{k+4} \theta_p N_p''(x))^2 dx.$$

It is immediate that the residual sum of squares about g can be rewritten

$$\sum_{i=1}^r \sum_{j=1}^{n_i} (y_{ij} - g(x_i))^2 = (\bar{Y} - F' \Theta)' D(\mu^n) (\bar{Y} - F' \Theta).$$

Using (2.7) we obtain

$$L(g) = (\bar{Y} - F'\Theta)'D(\mu^n)(\bar{Y} - F'\Theta) + \lambda\Theta'W\Theta. \quad (2.8)$$

Since λW is non-negative definite, the matrix $M(\mu^n) + \lambda W$ is strictly positive-definite. It therefore follows that (2.8) has a unique minimum, obtained by setting

$$g = \tilde{f}_n \quad \text{with} \quad \tilde{\Theta} = \Gamma^{-1}(\mu^n)F\mathcal{D}(\mu^n)\bar{Y}. \square$$

2.2 Smoothness Conditions

We will assume some regularity conditions on f . For $m \in \mathbb{N}$ and $M, L > 0$, let

$$\mathcal{F}_{m,M} = \{f \in \mathcal{C}^{m+1}(0,1) : \sup_{0 \leq x \leq 1} |f^{((m+1) \wedge 4)}(x)| \leq M\},$$

$$\mathcal{G}_{m,M,L} = \{f \in \mathcal{F}_{m,M} : \sup_{0 \leq x \leq 1} |f| \leq L\}.$$

The following lemma is an immediate consequence of theorem 2.59 of Schumaker(1981).

Lemma 2 For $f \in \mathcal{C}^m$ then,

$$\omega_m(f, t) \leq t^m \sup_{0 \leq x \leq 1} |f^{(m)}(x)| \quad \text{as} \quad t \rightarrow 0$$

where $\omega_m(f, \delta)$ denotes the m^{th} modulus of smoothness of f in L_∞ .

When $m=1$, the m^{th} modulus of smoothness is traditionally referred to as the modulus of continuity.

3 TEST OF NON-CONVEXITY

3.1 Convexity and Strict Convexity in $\mathcal{S}(k,4)$

Let us first characterize convexity in the class $\mathcal{S}(k,4)$. For a function g in the class $\mathcal{S}(k,4)$, we can write:

$$g(x) = \sum_{p=1}^{k+4} \theta_p N_p(x) \quad \text{with} \quad \Theta = (\theta_1, \dots, \theta_{k+4})' \in \mathbb{R}^{k+4}$$

$$\text{Then:} \quad g''(\eta_l) = \sum_{p=1}^{k+4} \theta_p N''_p(\eta_l) = \sum_{p=1}^{k+4} \theta_p d_{p,l},$$

where the coefficients $d_{p,l}$ are easily calculated from the knots (see Dierckx (1980))

$$\left\{ \begin{array}{l} d_{p,l} = 0 \quad \text{if } p \leq l \quad \text{or} \quad p \geq l + 4 \\ d_{l+1,l} = \frac{6}{(t_{l+5}-t_{l+2})(t_{l+5}-t_{l+3})} \\ d_{l+3,l} = \frac{6}{(t_{l+6}-t_{l+3})(t_{l+5}-t_{l+3})} \\ d_{l+2,l} = -(d_{l+3,l} + d_{l+1,l}) \end{array} \right. \quad \text{for } l = 0, \dots, k+1$$

Let $B_l = (0, 0, \dots, 0, -d_{l+1,l}, -d_{l+2,l}, -d_{l+3,l}, 0, \dots, 0)' \in \mathbb{R}^{k+4}$ and $\Theta = (\theta_1, \dots, \theta_{k+4})'$, then

$$g''(\eta_l) = -B'_l \Theta.$$

We have already mentioned in the introduction that g is a convex function if and only if $g''(\eta_l)$ is non-negative for all $l=0, \dots, k+1$ or equivalently $B'_l \Theta \leq 0$ for all $l=0, \dots, k+1$.

Characterizing similarly strictly convex functions does not seem possible but the following lemma gives a necessary condition.

Lemma 3 *Let $g(x) = \sum_{p=1}^{k+4} \theta_p N_p(x)$, let $C_l = B_l + B_{l+1}, l = 0, \dots, k$ then if g is strictly convex we have for all $l = 0, \dots, k$ $-C'_l \Theta > 0$.*

Proof

Let us assume that g is strictly convex. Since $g \in \mathcal{S}(k,4)$, then g'' is a linear function between any pair of adjacent knots η_i and η_{i+1} . Therefore, we can write

$$g''(x) = a_i x + b_i \quad \text{for all } x \in (\eta_i, \eta_{i+1}).$$

Moreover $g''(x) \geq 0$ or equivalently $a_i x + b_i \geq 0, i = 0, \dots, k$. Hence g is strictly convex if and only if

$$\left\{ \begin{array}{l} \text{if } a_i > 0 \quad \text{then } a_i \eta_i + b_i = g''(\eta_i) \geq 0 \\ \text{if } a_i < 0 \quad \text{then } a_i \eta_{i+1} + b_i = g''(\eta_{i+1}) \geq 0 \\ \text{if } a_i = 0 \quad \text{then } b_i > 0 \end{array} \right.$$

But $a_i = \frac{g''(\eta_{i+1}) - g''(\eta_i)}{\eta_{i+1} - \eta_i}$ and $b_i = g''(\eta_i) - \frac{g''(\eta_{i+1}) - g''(\eta_i)}{\eta_{i+1} - \eta_i} \eta_i$; therefore g is strictly convex if and only if

$$\left\{ \begin{array}{l} \text{if } g''(\eta_{i+1}) > g''(\eta_i) \quad \text{then } g''(\eta_i) \geq 0 \\ \text{if } g''(\eta_{i+1}) < g''(\eta_i) \quad \text{then } g''(\eta_{i+1}) \geq 0 \\ \text{if } g''(\eta_{i+1}) = g''(\eta_i) \quad \text{then } g''(\eta_i) > 0 \end{array} \right.$$

Hence, if g is strictly convex then $-(B'_l + B'_{l+1})\Theta > 0. \square$

It is easy to see that there are no redundant vectors in $\{C_0, \dots, C_k\}$ i.e there is no C_j such that

$$\{\Theta : -C'_l \Theta > 0, l = 0, \dots, k\} = \{\Theta : -C'_l \Theta > 0, l = 0, \dots, k, l \neq j\} \quad (3.1).$$

3.2 Definition of the Test

We intend to construct a test H_0 of: “ f is not strictly convex” against H_1 : “ f is strictly convex.” Lemma 3 does not characterize strict convexity, however we have the following property. Define \mathcal{C}_k to be the class of functions $g \in \mathcal{C}^2(0,1)$ satisfying the set of conditions $g''(\eta_l) + g''(\eta_{l+1}) > 0, l = 0, \dots, k$ associated to k knot points. If k tends to infinity, the union of all knot points defined by (2.1) is dense in $(0,1)$. It follows that a function g in $\mathcal{C}^2(0,1)$ belongs to $\bigcap_{k=1}^{\infty} \mathcal{C}_k$ if and only if g is strictly convex.

It is therefore natural to weaken H_1 replacing it by $H'_1 : f \in \mathcal{C}_k$, for some large k . It is then possible to construct a test for \bar{H}'_1 against H'_1 , using a likelihood ratio test (LRT) procedure studied by Berger (1989). Berger showed that if X is a p variate (in our case $p = k + 4$) normal random vector with unknown mean Θ and nonsingular covariance matrix Σ and if C_0, \dots, C_k satisfy (3.1), then the size- α LRT with null hypothesis “ $-C'_l\Theta \leq 0$ for some $l = 0, \dots, k$ ” versus the alternative hypothesis “ $-C'_l\Theta > 0$ for all $l = 0, \dots, k$ ” is the test which rejects the null hypothesis if

$$-\frac{C'_l X}{(C'_l \Sigma C_l)^{1/2}} \geq q_\alpha \quad \text{for all } l = 0, \dots, k.$$

where q_α is the upper 100α percentile of the standard normal distribution.

On the other hand, Beatson (1982) shows that for a smooth and convex function $f \in \mathcal{C}^m(0,1)$ ($0 \leq m \leq 3$), the uniform distance between f and the set $\mathcal{S}^{**}(k,4)$ of convex functions of $\mathcal{S}(k,4)$ tends to zero when the mesh size δ_k tends to zero (see lemma 6 below).

Therefore, for $f \in \mathcal{C}^m(0,1)$ ($0 \leq m \leq 3$), in order to test H_0 against H_1 since $\frac{\sigma^2}{n} M^{-1}(\mu^n)$ and $\frac{\sigma^2}{n} \Gamma^{-1}(\mu^n) M(\mu^n) \Gamma^{-1}(\mu^n)$ are respectively the covariance matrices of $\hat{\Theta}$ and $\tilde{\Theta}$, it is natural to define a test (T) by rejecting H_0 when

$$-\frac{\sqrt{n} C'_l \hat{\Theta}}{\sigma(C'_l M^{-1}(\mu^n) C_l)^{1/2}} \geq q_\alpha \quad \text{for all } l = 0, \dots, k$$

for the least squares estimator of f . Now, for the hybrid cubic spline estimator of f we define a test (T') by rejecting H_0 when

$$-\frac{\sqrt{n} C'_l \tilde{\Theta}}{\sigma(C'_l \Gamma^{-1}(\mu^n) M(\mu^n) \Gamma^{-1}(\mu^n) C_l)^{1/2}} \geq q_\alpha \quad \text{for all } l = 0, \dots, k.$$

3.3 Asymptotic Properties

To derive asymptotic properties of (T) and (T'), we assume next that μ^n converges to a design measure μ , where μ is an absolutely continuous measure. We denote by H_n and H the cumulative distribution function of μ^n and μ respectively. The

number of knots and the smoothing parameter will be functions of the sample size $k = k_n$ and $\lambda = \lambda_n$. The critical regions of (T) and (T') are respectively:

$$\Omega_n = \left\{ -\frac{\sqrt{n}C'_l\hat{\Theta}}{\sigma(C'_lM^{-1}(\mu^n)C_l)^{1/2}} \geq q_\alpha \quad \text{for all } l = 0, \dots, k_n \right\}$$

and

$$\Omega'_n = \left\{ -\frac{\sqrt{n}C'_l\tilde{\Theta}}{\sigma(C'_l\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)C_l)^{1/2}} \geq q_\alpha \quad \text{for all } l = 0, \dots, k_n \right\}.$$

In the following theorems, we give a result about the size and the power of (T) and we approximate the distribution of the statistic by a gaussian distribution with mean which converges to infinity. A measure of the distance between two distributions is given by the following modification of the Mallows distance (This measure was used by Hardle and Mammen (1993))

$$d(\mu, \nu) = \inf_{X, Y} (\mathbb{E} \|X - Y\|^2 \wedge 1 : \mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu).$$

Convergence in this metric is equivalent to weak convergence (see Hardle and Mammen (1993)).

Consider the following additional notations:

$$f''_\delta = \left(\frac{f''(\eta_0) + f''(\eta_1)}{\sigma(C'_0M^{-1}(\mu^n)C_0)^{1/2}}, \dots, \frac{f''(\eta_{k_n}) + f''(\eta_{k_n+1})}{\sigma(C'_{k_n}M^{-1}(\mu^n)C_{k_n})^{1/2}} \right)$$

and

$$f''_{\delta, \lambda} = \left(\frac{f''(\eta_0) + f''(\eta_1)}{\sigma(C'_0\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)C_0)^{1/2}}, \dots, \frac{f''(\eta_{k_n}) + f''(\eta_{k_n+1})}{\sigma(C'_{k_n}\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)C_{k_n})^{1/2}} \right).$$

Let

$$T_n^l(f) = -\frac{\sqrt{n}C'_l\hat{\Theta}}{\sigma(C'_lM^{-1}(\mu^n)C_l)^{1/2}} \quad l = 0, \dots, k_n$$

and

$$T_n(f) = (T_n^0(f), \dots, T_n^{k_n}(f))'.$$

$T_n^l(f)$ and $T'_n(f)$ are given by similar definitions.

Theorem 1 *Let $f \in \mathcal{F}_{m, M}$ with $m \geq 2$. Let us consider the hypothesis H_0 : “ f is not strictly convex” and H_1 : “ f is strictly convex.”*

Then, under the following assumptions

(i) ε_{ij} $i = 1, \dots, r ; j = 1, \dots, n_i$, *i.i.d* with mean zero and variance $\sigma^2 < +\infty$.

$$\begin{aligned}
& (ii) n_i \longrightarrow +\infty \quad i = 1, \dots, r . \\
& (iii) \sup_{0 \leq x \leq 1} |H_n(x) - H(x)| = o(k_n^{-1}), \quad \text{as } k_n \longrightarrow +\infty \\
& (iv) \lim_{n \rightarrow +\infty} r^{1/2} n^{1/2} \delta_{k_n}^{(m+1) \wedge 4} \left(\sup_{1 \leq i \leq r} \mu_i \right)^{1/2} = 0
\end{aligned}$$

the test (T) which rejects H_0 if

$$-\frac{\sqrt{n} C'_l \hat{\Theta}}{\sigma(C'_l M^{-1}(\mu^n) C_l)^{1/2}} \geq q_\alpha \quad \text{for all } l = 0, \dots, k_n$$

has asymptotically size α . More precisely,

$$\lim_{n \rightarrow +\infty} \sup_{f \in H_0} \mathcal{P}_f(\Omega_n) = \alpha.$$

We now give a result on the power of the test (T).

Theorem 2 *Let us assume that $m \geq 2$. Under the assumptions of theorem 1 if*

$$(v) \lim_{n \rightarrow +\infty} n^{1/2} \delta_{k_n}^{5/2} = +\infty$$

then for $f \in H_1$ and $f'' > 0$

$$\lim_{n \rightarrow +\infty} \mathcal{P}_f(\Omega_n) = 1.$$

Now it would be interesting to know how the test behaves under the local alternatives. We consider a sequence of a local alternatives:

$$f_n(x) = g(x) + h_n \gamma(x) \quad (*)$$

where the know function $\gamma(\cdot)$ lies in $\mathcal{F}_{m,M}$ and g lies in the null hypothesis. The following theorem gives the power of our test against local misspecification (*).

Theorem 3 *Let us assume that $m \geq 3$. Under the assumptions (i), (iii) and (v) if*

$$\begin{aligned}
& \varepsilon_{ij} \rightsquigarrow \mathcal{N}(0, \sigma^2), \\
& \lim_{n \rightarrow +\infty} r^{1/2} n^{1/2} \delta_{k_n}^3 \left(\sup_{1 \leq i \leq r} \mu_i \right)^{1/2} = 0
\end{aligned}$$

and

$$h_n = (n\delta_{k_n}^5)^{-1/2}$$

then

$$d(\mathcal{L}(T_n(g + h_n\gamma)), \mathcal{N}(\sqrt{n}g'' + \delta_{k_n}^{-5/2}\gamma''_s, V(\mu^n))) \rightarrow 0.$$

where

$$V(\mu^n) = C'M^{-1}(\mu^n)C$$

and

$$C = \left(\frac{C_0}{(C'_0 M^{-1}(\mu^n) C_0)^{1/2}}, \dots, \frac{C_{k_n}}{(C'_{k_n} M^{-1}(\mu^n) C_{k_n})^{1/2}} \right).$$

We have similar theorems for the test (T') in the smooth class of functions $\mathcal{G}_{m,M,L}$.

Theorem 4 Let $f \in \mathcal{G}_{m,M,L}$ with $m \geq 2$.

Under the assumptions of theorem 1 and if

$$(iv') \lim_{n \rightarrow +\infty} r^{1/2} n^{1/2} \delta_{k_n}^{-7/2} \lambda_n(\sup_{1 \leq i \leq r} \mu_i)^{1/2} = 0,$$

then the test (T') has asymptotically size smaller than α .

Theorem 5 Under the assumptions of theorems 2 and 4 we have for all f in H_1 such that $f'' > 0$

$$\lim_{n \rightarrow +\infty} \mathcal{P}_f(\Omega'_n) = 1.$$

Theorem 6 Under the assumptions of theorems 3 and 4, if

$$\lim_{n \rightarrow +\infty} r^{1/2} \delta_{k_n}^{-13/2} \lambda_n(\sup_{1 \leq i \leq r} \mu_i)^{1/2} = 0$$

then

$$d(\mathcal{L}(T'_n(g + h_n\gamma)), \mathcal{N}(\sqrt{n}(g'' + \delta_{k_n}^{-5/2}\gamma''_s), V^1(\mu^n))) \rightarrow 0.$$

where

$$V^1(\mu^n) = B'\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)B$$

and

$$B = \left(\frac{C_0}{(C'_0 \Gamma^{-1}(\mu^n) M(\mu^n) \Gamma^{-1}(\mu^n) C_0)^{1/2}}, \dots, \frac{C_{k_n}}{(C'_{k_n} \Gamma^{-1}(\mu^n) M(\mu^n) \Gamma^{-1}(\mu^n) C_{k_n})^{1/2}} \right).$$

Remarks

- Assumption (iii) is the same as in Aggarwal and Studden (1980)
- It is easy to see that assumption (iv) implies that

$$\lim_{n \rightarrow +\infty} k_n = +\infty$$

- For a uniform design, i.e $\mu_i = 1/r, i = 1, \dots, r$ then (iv) becomes

$$\lim_{n \rightarrow +\infty} n^{1/2} \delta_{k_n}^{(m+1)\wedge 4} = 0$$

This is the case for example when there are n distinct points for n measurements ($r=n$).

- If in particular λ_n satisfies

$$\lim_{n \rightarrow +\infty} \lambda_n \delta_{k_n}^{-7/2-(m+1)\wedge 4} = 0,$$

then (iv') is satisfied.

- Finally, note that the test procedure requires the knowledge of the variance σ^2 . However, the results in this article can be considered approximately true if σ is unknown but, in order to compute the statistic, we need a consistent estimate of σ . This can be obtained in the case of the least squares estimator, using $\hat{\sigma}_n = \frac{1}{n-(k+4)} \sum_{i=1}^r (\bar{y}_i - \hat{f}_n(x_i))^2$ where \bar{y}_i and \hat{f}_n are defined by (2.3) and (2.4) and $\tilde{\sigma}_{n,\lambda_n}$ (for a fixed λ_n) has a similarly definition or alternatively, any consistent estimator based on nonparametric regression techniques.

For the proof of theorems 1 to 6 we need some preliminaries lemmas. Lemma 4 from Schumaker (1981) gives a sup norm error bound when approximating a function in $\mathcal{C}^m(0, 1)$ ($0 \leq m \leq 3$) with a cubic spline.

Lemma 4 *Let $0 \leq m \leq 3$. There is a constant c such that for all $f \in \mathcal{C}^m(0, 1)$, there is a function S in $\mathcal{S}(k_n, 4)$ such that:*

$$\sup_{0 \leq x \leq 1} |f^{(r)}(x) - S^{(r)}(x)| \leq c \delta_{k_n}^{m-r} \omega_{4-m}(f^{(m)}, \delta_{k_n}) \quad \text{for all } r = 0, \dots, m.$$

Using this property and lemma 2, we easily get the following:

Lemma 5 *Let $m \geq 0$. There is a constant c such that for all $f \in \mathcal{F}_{m,M}$, there is a function S in $\mathcal{S}(k_n, 4)$ such that:*

$$\sup_{0 \leq x \leq 1} |f^{(r)}(x) - S^{(r)}(x)| \leq c \delta_{k_n}^{((m+1)\wedge 4)-r} \quad \text{for all } r = 0, \dots, m \wedge 3.$$

Lemma 6 from Beatson (1982) gives a similar sup norm error bound when approximating a convex function of $\mathcal{C}^m(0, 1)$ ($0 \leq m \leq 3$) with a convex cubic spline.

Lemma 6 *There is a constant c such that if $f \in \mathcal{C}^m(0, 1)$ ($0 \leq m \leq 3$) is convex, then there is a convex function S in $\mathcal{S}(k_n, 4)$ such that :*

$$\sup_{0 \leq x \leq 1} |f(x) - S(x)| \leq c \delta_{k_n}^m \omega(f^{(m)}, \delta_{k_n}).$$

The following lemma is obtained by a straightforward manipulation of results of Agarwal and Studden (1980).

Lemma 7 *Under the assumptions of theorem 1,*

$$1) \| M^{-1}(\mu^n) \| = \| (F\mathcal{D}(\mu^n)F')^{-1} \| = \mathcal{O}(\delta_{k_n}^{-1}), \quad \text{as } n \rightarrow +\infty$$

$$2) \| M(\mu^n) \| = \mathcal{O}(\delta_{k_n}), \quad \text{as } n \rightarrow +\infty$$

where $\| A \| =_{def} \max_x (\| Ax \| / \| x \|)$ with $\| x \| = (x'x)^{1/2}$.

Proof: Equality 1) follows from lemma 6.6 of Agarwal and Studden (1980). Now, recall that

$$M(\mu^n) = \sum_{i=1}^r N(x_i)N'(x_i)\mu_i.$$

We can write:

$$M(\mu^n) = \int N(x)N'(x)d\mu^n(x).$$

Let

$$M(\mu) = \int N(x)N'(x)d\mu(x).$$

Agarwal and Studden (1980) prove that $M(\mu) = \mathcal{O}(\delta_{k_n})$ as $n \rightarrow +\infty$ in lemma 6.3.

As in equality 6.22 of Agarwal and Studden (1980) pp. 1319, we can write

$$M^{-1}(\mu^n) = [I - U_n]^{-1}M^{-1}(\mu)$$

where I is the $(k+4) \times (k+4)$ identity matrix and a U_n matrix such that $\| U_n \| = o(1)$. Hence

$$\| M(\mu^n) \| \leq \| M(\mu) \| \| I - U_n \|.$$

Now, since $M(\mu) = \mathcal{O}(\delta_{k_n})$, $n \rightarrow +\infty$ we have the result. \square

The following important result will be used in the proof of theorems 4 to 6.

Lemma 8 *Under the assumptions of theorem 3*

$$\| M^{-1}(\mu^n)W \| = \mathcal{O}(\delta_{k_n}^{-4}), \quad \text{as } n \rightarrow +\infty.$$

Proof: We have

$$\| M^{-1}(\mu^n)W \| \leq \| M^{-1} \| \| W \|.$$

Using a classical inequality (see Ciarlet and Thomas (1982)) we obtain

$$\| W \| \leq \| W \|_\infty$$

where

$$\| A \|_\infty =_{def} \max_i \sum_j | a_{ij} |.$$

Hence, since $\| M^{-1} \| = \mathcal{O}(\delta_{k_n}^{-1})$ the lemma follows if we prove that

$$\| W \|_\infty = \mathcal{O}(\delta_{k_n}^{-3}).$$

In fact, because W is a band matrix, it suffice to prove that $| W_{ij} | = \mathcal{O}(\delta_{k_n}^{-3})$. Using formula (6.19) (page 1318) of Argarwal and Studden (1980) we see that for all $i = 1, \dots, k_n + 4$ and for $j = 1, 2$

$$\sup_{x \in (0,1)} | N_i^{(j)}(x) | = \mathcal{O}(k_n^j)$$

and since the sequence of knots is quasi-uniforme, we obtain

$$\sup_{x \in (0,1)} | N_i^{(j)}(x) | = \mathcal{O}(\delta_{k_n}^{-j}). \quad (**)$$

On the other hand

$$W_{ij} = \int_0^1 N_i''(x) N_j''(x) dx = N_i'(1) N_j''(1) - N_i'(0) N_j''(0) - \int_0^1 N_i'(x) N_j'''(x) dx.$$

Hence, using (**) we see that

$$| N_i'(1) N_j''(1) - N_i'(0) N_j''(0) | = \mathcal{O}(\delta_{k_n}^{-3}).$$

Now, we can write

$$\int_0^1 N_i'(x) N_j'''(x) dx = \sum_{p=0}^{p=k} \int_{\eta_p}^{\eta_{p+1}} N_i'(x) N_j'''(x) dx$$

and using the proof of lemma 3, we obtain (since for $x \in (\eta_p, \eta_{p+1})$, $N_j'''(x) = \frac{N_j''(\eta_{p+1}) - N_j''(\eta_p)}{\eta_{p+1} - \eta_p}$)

$$\left| \int_{\eta_p}^{\eta_{p+1}} N_i'(x) N_j'''(x) dx \right| \leq \sup_{x \in (0,1)} | N_i'(x) | \| N_j''(\eta_{p+1}) - N_j''(\eta_p) |.$$

Using notation of section (3.1), we can write $N_j''(\eta_{p+1}) = d_{j,p+1}$ and $N_j''(\eta_p) = d_{j,p}$. Hence the lemma follows from the (**) and the fact that $d_{j,p} = 0$ if $| p - j | \geq 4$ else $d_{j,p} = \mathcal{O}(\delta_{k_n}^{-2})$. \square

Proof of Theorem 1

Recall that

$$\widehat{\Theta} = M^{-1}(\mu^n)F\mathcal{D}(\mu^n)\bar{Y} = M^{-1}(\mu^n)F\mathcal{D}(\mu^n)(f_\Delta + \bar{\varepsilon})$$

hence, $\widehat{\Theta}$ has the same asymptotic distribution than

$$\mathcal{N}(M^{-1}(\mu^n)F\mathcal{D}(\mu^n)f_\Delta, \frac{\sigma^2}{n}M^{-1}(\mu^n)).$$

In what follows, we denoted this by

$$\widehat{\Theta} \rightsquigarrow \mathcal{N}_{As}(M^{-1}(\mu^n)F\mathcal{D}(\mu^n)f_\Delta, \frac{\sigma^2}{n}M^{-1}(\mu^n)).$$

Let $f \in \mathcal{F}_{m,M}$ ($m \geq 1$). By Lemma 5, there is a S in $\mathcal{S}(k_n, 4)$, such that :

$$\sup_{0 \leq x \leq 1} |f^{(r)}(x) - S^{(r)}(x)| \leq c\delta_{k_n}^{((m+1) \wedge 4) - r} \quad \text{for all } r = 0, \dots, m \wedge 3. \quad (3.2)$$

$S \in \mathcal{S}(k_n, 4)$ hence, there exists $\Theta^0 \in \mathbb{R}^{k+4}$ such that $S(x) = N'(x)\Theta^0$.

Let

$$S_\Delta = (S(x_1), \dots, S(x_r))' = (N'(x_1)\Theta^0, \dots, N'(x_r)\Theta^0)' = F'\Theta^0.$$

Then,

$$M^{-1}(\mu^n)F\mathcal{D}(\mu^n)S_\Delta = M^{-1}(\mu^n)F\mathcal{D}(\mu^n)F'\Theta^0 = \Theta^0.$$

Hence

$$E\widehat{\Theta} = \Theta^0 + M^{-1}(\mu^n)F\mathcal{D}(\mu^n)(f_\Delta - S_\Delta) = \Theta^0 + b_n \quad \text{say}$$

Now, let us assume that f is not strictly convex, then there exists an interval $I \subseteq (0, 1)$ such that for all $x \in I$, $f''(x) \leq 0$ (f is concave in I).

Therefore there exists l_n with $0 \leq l_n \leq k_n$ such that

$$C'_{l_n}\Theta^0 = -(S''(\eta_{l_n}) + S''(\eta_{l_n+1})) \geq 0. \quad (3.3)$$

Indeed, if for all $x \in I$, $f''(x) < 0$, then using (3.2) we see that for n sufficiently large, there exists an interval $J \subseteq I$ such that for all $x \in J$, $S''(x) \leq 0$ and we can take η_{l_n} and $\eta_{l_n+1} \in J$.

Now, if for all $x \in I$, $f''(x) = 0$ then, using formula 6.44 of theorem 6.20 in Schumaker (1981) we see that there is an interval $J \subseteq I$ such that for all $x \in J$, $S''(x) = 0$.

Now, let $R_l = \left\{ -\frac{\sqrt{n}C'_{l_n}\widehat{\Theta}}{\sigma(C'_{l_n}M^{-1}(\mu^n)C_{l_n})^{1/2}} \geq q_\alpha \right\}$ then, $\Omega_n = \bigcap_{l=0}^{k_n} R_l$ and therefore,

$$\mathcal{P}_f(\Omega_n) \leq \inf_{0 \leq l \leq k_n} \mathcal{P}_f(R_l) \leq \mathcal{P}_f(R_{l_n}).$$

Hence,

$$\begin{aligned} \mathcal{P}_f(\Omega_n) &\leq \mathcal{P}_f\left(\frac{-\sqrt{n}C'_{l_n}\hat{\Theta}}{\sigma(C'_{l_n}M^{-1}(\mu^n)C_{l_n})^{1/2}} \geq q_\alpha\right) \\ &\leq \mathcal{P}_f(\mathcal{N}_{As}(0,1) \geq q_\alpha + \frac{\sqrt{n}C'_{l_n}\Theta^0}{\sigma(C'_{l_n}M^{-1}(\mu^n)C_{l_n})^{1/2}} + \frac{\sqrt{n}C'_{l_n}b_n}{\sigma(C'_{l_n}M^{-1}(\mu^n)C_{l_n})^{1/2}}). \end{aligned}$$

Then, by (3.3) we have

$$\mathcal{P}_f(\Omega_n) \leq \mathcal{P}_f(\mathcal{N}_{As}(0,1) \geq q_\alpha + \frac{\sqrt{n}C'_{l_n}b_n}{\sigma(C'_{l_n}M^{-1}(\mu^n)C_{l_n})^{1/2}})$$

The proof consists of bounding the absolute value of $\frac{\sqrt{n}C'_{l_n}b_n}{\sigma(C'_{l_n}M^{-1}(\mu^n)C_{l_n})^{1/2}}$. We have $|C'_{l_n}b_n| \leq \|C_{l_n}\| \|b_n\|$ and

$$\|b_n\| \leq \|M^{-1}(\mu^n)\| \|F\mathcal{D}(\mu^n)\| \|f_\Delta - S_\Delta\|.$$

Since

$$\|F\mathcal{D}^{1/2}(\mu^n)\|^2 = \|F\mathcal{D}(\mu^n)F'\|,$$

we have:

$$\begin{aligned} \|F\mathcal{D}(\mu^n)\| &\leq \|F\mathcal{D}^{1/2}(\mu^n)\| \|\mathcal{D}^{1/2}(\mu^n)\| \leq \|M(\mu^n)\|^{1/2} \left(\sup_{1 \leq i \leq r} \mu_i\right)^{1/2} \\ \|f_\Delta - S_\Delta\| &= \left(\sum_{i=1}^r (f(x_i) - S(x_i))^2\right)^{1/2} \\ &\leq r^{1/2} \sup_{0 \leq x \leq 1} |f(x) - S(x)| \leq r^{1/2} c\delta_{k_n}^{(m+1)\wedge 4}. \end{aligned}$$

Now,

$$\begin{aligned} (C'_{l_n}M^{-1}(\mu^n)C_{l_n})/\|C_{l_n}\|^2 &\geq \min_X (X'M^{-1}(\mu^n)X)/\|X\|^2 = (1/\|M^{1/2}(\mu^n)\|)^2 \\ &\geq 1/\|M(\mu^n)\|. \end{aligned}$$

Hence, $\|C_{l_n}\|/(C'_{l_n}M^{-1}(\mu^n)C_{l_n})^{1/2} \leq \|M(\mu^n)\|^{1/2}$. Therefore,

$$\|C_{l_n}\| \|b_n\| / (C'_{l_n}M^{-1}(\mu^n)C_{l_n})^{1/2} \leq \|b_n\| \|M(\mu^n)\|^{1/2}. \quad (3.4)$$

Hence

$$\sup_{f \in H_0} \frac{\sqrt{n} |C'_{l_n}b_n|}{\sigma(C'_{l_n}M^{-1}(\mu^n)C_{l_n})^{1/2}} \leq \frac{\sqrt{n}}{\sigma} r^{1/2} c_0 \delta_{k_n}^{(m+1)\wedge 4} \left(\sup_{1 \leq i \leq r} \mu_i\right)^{1/2} \quad (3.5)$$

for some constant c_0 .

Therefore, using (3.5) we have

$$\sup_{f \in H_0} \mathcal{P}_f(\Omega_n) \leq \mathcal{P}(\mathcal{N}_{As}(0, 1)) \geq q_\alpha - \frac{\sqrt{n}}{\sigma} r^{1/2} c_0 \delta_{k_n}^{(m+1) \wedge 4} \left(\sup_{1 \leq i \leq r} \mu_i \right)^{1/2}.$$

Hence, by assumption (iii)

$$\lim_{n \rightarrow +\infty} \sup_{f \in H_0} \mathcal{P}_f(\Omega_n) \leq \alpha.$$

On the other hand,

$$\sup_{f \in H_0} \mathcal{P}_f(\Omega_n) \geq \sup_{f \in H_0 \cap \mathcal{S}(k_n, 4)} \mathcal{P}_f(\Omega_n).$$

But, if $f \in H_0 \cap \mathcal{S}(k_n, 4)$, there is $\Theta \in \mathbb{R}^{k_n+4}$ such that

$$f(x) = \Theta' N(x).$$

Hence $f \in H_0 \cap \mathcal{S}(k_n, 4)$ if and only if $\Theta \in \mathcal{K}$

where

$$\mathcal{K} = \{x \in \mathbb{R}^{k_n+4} : \exists l, \quad 0 \leq l \leq k_n : -C'_l x \leq 0\}.$$

Berger (1989) prove that

$$\sup_{\nu \in \mathcal{K}, \nu = \mathbb{E}\hat{\Theta}} \mathcal{P}_\nu(\Omega_n) = \alpha$$

therefore

$$\lim_{n \rightarrow +\infty} \sup_{f \in H_0} \mathcal{P}_f(\Omega_n) = \alpha. \square$$

Proof of Theorem 2

We want to show that for all f in H_1 with $f'' > 0$ that

$$\lim_{n \rightarrow +\infty} \mathcal{P}_f(\Omega_n) = 1.$$

Let $f \in \mathcal{F}_{m,M}$ ($m \geq 2$). Let us assume that f is strictly convex, then by lemma 6, there is a S in $\mathcal{S}(k_n, 4)$ convex such that :

$$\sup_{0 \leq x \leq 1} |f(x) - S(x)| \leq c \delta_{k_n}^{(m+1) \wedge 4}, \quad (3.6).$$

As above in (3.5), it is easy to see that

$$\left| \frac{\sqrt{n} C'_l M^{-1}(\mu^n) F \mathcal{D}(\mu^n) (f_\Delta - S_\Delta)}{\sigma (C'_l M^{-1}(\mu^n) C_l)^{1/2}} \right| \leq \frac{\sqrt{n}}{\sigma} r^{1/2} c_1 \delta_{k_n}^{(m+1) \wedge 4} \left(\sup_{1 \leq i \leq r} \mu_i \right)^{1/2}.$$

Therefore, under the assumption (iv), we see that:

$$\lim_{n \rightarrow +\infty} \sup \frac{\sqrt{n} C'_l M^{-1}(\mu^n) F \mathcal{D}(\mu^n)(f_\Delta - S_\Delta)}{\sigma(C'_l M^{-1}(\mu^n) C_l)^{1/2}} = 0 \quad (3.7)$$

On the other hand, for all $l = 0, \dots, k_n$.

$$\frac{1}{\|M(\mu^n)\|} \leq C'_l M^{-1}(\mu^n) C_l / \|C_l\|^2 \leq \|M^{-1}(\mu^n)\|.$$

Hence,

$$\frac{\|M(\mu^n)\|^{1/2}}{\|C_l\|} \geq \frac{1}{(C'_l M^{-1}(\mu^n) C_l)^{1/2}} \geq \frac{1}{\|C_l\| \|M^{-1}(\mu^n)\|^{1/2}}. \quad (3.8)$$

Now, since the sequence of knots is quasi-uniforme (see (2.2)), it is easy to see that there are constants $a_1 > 0$ and $a_2 > 0$ such that

$$a_1 \delta_{k_n}^{-2} \leq \|C_l\| \leq a_2 \delta_{k_n}^{-2}.$$

Hence

$$\frac{\sqrt{n} C'_l \Theta^0}{\sigma(C'_l M^{-1}(\mu^n) C_l)^{1/2}} = -\frac{\sqrt{n}(S''(\eta_l) + S''(\eta_{l+1}))}{\sigma(C'_l M^{-1}(\mu^n) C_l)^{1/2}} \leq -a \sqrt{n} \delta_{k_n}^{5/2} (S''(\eta_l) + S''(\eta_{l+1})) \quad (3.9)$$

for some constant $a > 0$. Now,

$$\mathcal{P}_f(\Omega_n) = \mathcal{P}_f\left(\bigcap_{0 \leq l \leq k_n} R_l\right).$$

If f is strictly convex we see that by (3.2) S is convex for n sufficiently large. Hence $C'_l \Theta^0 \leq 0$ for all $l = 0, \dots, k_n$. Therefore, using (3.9) we see that

$$\frac{\sqrt{n} C'_l \Theta^0}{\sigma(C'_l M^{-1}(\mu^n) C_l)^{1/2}} \leq -a \sqrt{n} \delta_{k_n}^{5/2} \inf_{0 \leq x \leq 1} f''(x).$$

Hence,

$$\{\mathcal{N}_{As}(0, 1) \geq q_\alpha - c_4 \sqrt{n} \delta_{k_n}^{5/2} + \frac{\sqrt{n}}{\sigma} r^{1/2} c_0 \delta_{k_n}^{(m+1) \wedge 4} (\sup_{1 \leq i \leq r} \mu_i)^{1/2}\} \subseteq R_l \quad \text{for all } l = 0, \dots, k_n$$

where $c_3 = a \inf_{0 \leq x \leq 1} f''(x)$. Moreover, under the assumptions (iv) and (v) we see that

$$\lim_{n \rightarrow +\infty} \mathcal{P}_f(\mathcal{N}_{As}(0, 1) \geq q_\alpha - c_3 \sqrt{n} \delta_{k_n}^{5/2} + \frac{\sqrt{n}}{\sigma} r^{1/2} c_0 \delta_{k_n}^{(m+1) \wedge 4} (\sup_{1 \leq i \leq r} \mu_i)^{1/2}) = 1.$$

Hence Hence for all f in H_1 and such that $f'' > 0$ then

$$\lim_{n \rightarrow +\infty} \mathcal{P}_f(\Omega_n) = 1. \quad \square$$

Proof of Theorems 3

We have $f_n = g + h_n \gamma$ therefore we can write

$$\hat{\Theta} = M^{-1}(\mu^n)F\mathcal{D}(\mu^n)(g_\Delta + \bar{\varepsilon}) + h_n M^{-1}(\mu^n)F\mathcal{D}(\mu^n)\gamma_\Delta.$$

Now, we proceed as in the proof of theorem 1. For $l = 0, \dots, k_n$ we can write

$$\begin{aligned} \mathbb{E}T_n^l(g + h_n \gamma) &= -\frac{\sqrt{n}C'_l b_n}{\sigma(C'_l M^{-1}(\mu^n)C_l)^{1/2}} + \frac{\sqrt{n}(S''(\eta_l) + S''(\eta_{l+1}))}{\sigma(C'_l M^{-1}(\mu^n)C_l)^{1/2}} \\ &\quad - \frac{h_n \sqrt{n}C'_l M^{-1}(\mu^n)F\mathcal{D}(\mu^n)\gamma_\Delta}{\sigma(C'_l M^{-1}(\mu^n)C_l)^{1/2}}. \end{aligned}$$

Arguing as above in (3.5) and using (3.2) and (3.8) we can see that

$$\begin{aligned} & \left| \mathbb{E}T_n^l(g + h_n \gamma) - \frac{\sqrt{n}(f''(\eta_l) + f''(\eta_{l+1}))}{\sigma(C'_l M^{-1}(\mu^n)C_l)^{1/2}} + h_n \frac{\sqrt{n}(\gamma''(\eta_l) + \gamma''(\eta_{l+1}))}{\sigma(C'_l M^{-1}(\mu^n)C_l)^{1/2}} \right| \\ &= \mathcal{O}(\sqrt{n}\delta_{k_n}^{(m+3/2)\wedge 9/2}) + \mathcal{O}(\sqrt{n}r^{1/2}\delta_{k_n}^{(m+1)\wedge 4}(\sup_{1 \leq i \leq r} \mu_i)^{1/2}) \\ &\quad + \mathcal{O}(r^{1/2}\delta_{k_n}^{3/2}(\sup_{1 \leq i \leq r} \mu_i)^{1/2} + \delta_{k_n}^2). \end{aligned}$$

Now, let

$$Y_n = T_n(g + h_n \gamma) - \mathbb{E}T_n(g + h_n \gamma) + \sqrt{n}f''_\delta + \delta_{k_n}^{-5/2}\gamma''_\delta$$

Under the assumptions of theorem 3, we see that

$$Y_n \rightsquigarrow \mathcal{N}(\sqrt{n}f''_\delta + \delta_{k_n}^{-5/2}\gamma''_\delta, V(\mu^n))$$

and

$$\mathbb{E} \| Y_n - T_n(g + h_n \gamma) \|^2 \rightarrow 0.$$

hence

$$d(\mathcal{L}(T_n(g + h_n \gamma)), \mathcal{N}(\sqrt{n}f''_\delta + \delta_{k_n}^{-5/2}\gamma''_\delta, V(\mu^n))) \rightarrow 0. \quad \square$$

Proof of Theorems 4 to 6

Let us recall that

$$\tilde{\Theta} = \Gamma^{-1}(\mu^n)F\mathcal{D}(\mu^n)\bar{Y} = (M(\mu^n) + \lambda_n W)^{-1}F\mathcal{D}(\mu^n)\bar{Y}.$$

Therefore

$$\mathbb{E}\tilde{\Theta} = (M(\mu^n) + \lambda_n W)^{-1}F\mathcal{D}(\mu^n)f_\Delta.$$

$$= \{(M(\mu^n) + \lambda_n GQR^{-1}Q'G')^{-1}F\mathcal{D}(\mu^n)f_\Delta - M^{-1}(\mu^n)F\mathcal{D}(\mu^n)f_\Delta\} + IE\hat{\Theta}.$$

Let

$$a_n = (M(\mu^n) + \lambda_n W)^{-1}F\mathcal{D}(\mu^n)f_\Delta - M^{-1}(\mu^n)F\mathcal{D}(\mu^n)f_\Delta.$$

Hence, as above in the proof of theorem 1, we can write

$$IE\tilde{\Theta} = a_n + b_n + \Theta^0.$$

Then, for all $f \in \mathcal{G}_{m,M,L}$ we have

$$\begin{aligned} \mathcal{P}_f(\Omega'_n) &= \mathcal{P}_f(\mathcal{N}(0, 1)) \geq q_\alpha + \frac{\sqrt{n}C'_l\Theta^0}{\sigma(C'_l\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)C_l)^{1/2}} + \\ &\frac{\sqrt{n}C'_lb_n}{\sigma(C'_l\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)C_l)^{1/2}} + \frac{\sqrt{n}C'_la_n}{\sigma(C'_l\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)C_l)^{1/2}}. \end{aligned}$$

Therefore as in the proofs of theorems 1 to 3, theorems 4, 5 and 6 will be established when we show that

$$\lim_{n \rightarrow +\infty} \sup_{f \in H_0} \frac{\sqrt{n} |C'_lb_n|}{\sigma(C'_l\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)C_l)^{1/2}} = 0, \quad (3.10)$$

and

$$\lim_{n \rightarrow +\infty} \sup_{f \in H_0} \frac{\sqrt{n} |C'_la_n|}{\sigma(C'_l\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)C_l)^{1/2}} = 0. \quad (3.11)$$

As above in (3.4) it can be shown that

$$\begin{aligned} &\|C_l\| \|b_n\| / (C'_l\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)C_l)^{1/2} \\ &\leq \|b_n\| \|\Gamma(\mu^n)\| \|M^{-1}(\mu^n)\|^{1/2}, \quad (3.12) \end{aligned}$$

and

$$\begin{aligned} &\|C_l\| \|a_n\| / (C'_l\Gamma^{-1}(\mu^n)M(\mu^n)\Gamma^{-1}(\mu^n)C_l)^{1/2} \\ &\leq \|a_n\| \|\Gamma(\mu^n)\| \|M^{-1}(\mu^n)\|^{1/2}. \quad (3.13) \end{aligned}$$

Therefore, the proof consist of bounding $\|\Gamma(\mu^n)\|$ and $\|a_n\|$. Under the assumptions of theorem 4 using lemma 8 we see that $\|\lambda_n M^{-1}(\mu^n)W\| < 1$ for n sufficiently large. We can then invert $(I + \lambda_n M^{-1}(\mu^n)W)$ using a power series expansion,

$$\begin{aligned} (I + \lambda_n M^{-1}(\mu^n)W)^{-1} &= I - \lambda_n M^{-1}(\mu^n)W + (\lambda_n M^{-1}(\mu^n)W)^2 + \dots \\ &= I + U \quad \text{say.} \end{aligned}$$

Using the matrix norm properties, we can show that

$$\| U \| \leq \| \lambda_n M^{-1}(\mu^n) W \| / (1 - \| \lambda_n M^{-1}(\mu^n) W \|) = \mathcal{O}(\lambda_n \delta_{k_n}^{-4}) \quad \text{by lemma 8.}$$

We have

$$\Gamma^{-1}(\mu^n) = (I + U)M^{-1}(\mu^n),$$

and therefore

$$\Gamma(\mu^n) = M(\mu^n)(I + U)^{-1}.$$

Standard arguments (e.g. Ciarlet and Thomas (1982)) show that

$$\| (I + U)^{-1} \| \leq 1 / (1 - \| U \|).$$

Hence

$$\| \Gamma(\mu^n) \| \leq c \| M(\mu^n) \|, \quad (3.14)$$

for some constant c . Now, using (3.12) we see that there is a constant c_1 such that

$$\| C_l \| \| \| b_n \| \| / (C'_l \Gamma^{-1}(\mu^n) M(\mu^n) \Gamma^{-1}(\mu^n) C_l)^{1/2} \leq c_1 \| b_n \| \| M(\mu^n) \|^{1/2},$$

and (3.10) follows from the assumptions (iv) of theorem 1. Now, since $\Gamma^{-1}(\mu^n) = (I + U)M^{-1}(\mu^n)$, we see that $a_n = UM^{-1}(\mu^n)F\mathcal{D}(\mu^n)f_\Delta$. Hence it can be shown as in the proof of theorem 1

$$\| a_n \| \leq \| U \| \| M^{-1}(\mu^n) \| \| F\mathcal{D}(\mu^n) \| \| f_\Delta \| = \mathcal{O}(\lambda_n \delta_{k_n}^{-4} r^{1/2} L(\sup_{1 \leq i \leq r} \mu_i)^{1/2}).$$

Now, (3.11) follows from (3.13), (3.14) and assumption (iv') of theorem 4. \square

4 APPLICATION TO A REAL DATA SET

We have experimented our method on a real data set obtained from the investigation “*Formation Qualification Professionnelle*” realized by INSEE in 1993. Simioni (1995) used this data set in a nonparametric regression model to estimate the logarithm of wage as a function of a set of characteristic variables such as age in the firm, experience, ... Human capital theory predicts the concavity of the logarithm of wage as a function of the experience (see Berndt (1991)). In this data set, this curve displays however a low curvature and it is therefore interesting to test its concavity.

The sample size i.e the number of wage-earners is very large $n = 7948$. The deterministic points $x_i, i = 1, \dots, r$ are represented by the experience (the number of year since the first job). Table (4.1) from Simioni (1995) gives some characteristics on the empirical distribution of the variables wage and experience.

A least squares and a hybrid cubic splines estimators are fitted to the data with k_n equidistant knots, $\lambda_n = 0$ corresponds to the least squares estimator. The number k_n takes the values 3, 9, 15, 21, 27 and k_{n,λ_n}^{gcv} , where k_{n,λ_n}^{gcv} is the number of knots selected by cross-validation (when k_n varies in $\{3, 4, \dots, 27\}$) for fixed λ_n . Computations were done for a smoothing parameter λ_n equal to $0, 5 \cdot 10^{-3}, 5 \cdot 10^{-3.5}, 5 \cdot 10^{-4}, 5 \cdot 10^{-4.5}, 5 \cdot 10^{-5}$. Figure (1) shows the data set with the estimate corresponding to $\lambda_n = 5 \cdot 10^{-4}$ and $k_n = k_{n,\lambda_n}^{gcv} = 4$. To compute tests statistics (T) and (T'), we used respectively consistent estimates $\hat{\sigma}_n$ and $\tilde{\sigma}_{n,\lambda_n}$ of σ as in section 3.3. Table (4.2) displays a "1" when the test rejects H_0 .

The results of this table show that the test confirms economic theory for choices of the smoothing parameters k_n and λ_n corresponding to smooth estimates (i.e. for λ_n small enough and k_n not too large). This is not surprising since rough estimates will produce artificial changes in curvature. Therefore a good choice of these parameters is essential. In practice, one could choose simultaneously the two parameters k_n and λ_n .

5 CONCLUSION

A small simulation study in Diack (1996) shows that the finite sample behavior of the test is quite satisfactory. It also reinforces our belief that the hybrid spline test is more reliable.

A test of non-monotonicity can be readily constructed paralleling the above non-convexity test with quadratic splines instead of cubic splines. Indeed, Beatson (1982) also shows that for a monotone function $f \in \mathcal{C}^m(0, 1)$ ($1 \leq m \leq 2$), the uniform distance between f and the set $\mathcal{S}^*(k, 3)$ of monotones functions of $\mathcal{S}(k, 3)$ tends to zero when the mesh size δ_{k_n} tends to zero.

It would be nice to extend this framework to multivariate regression models. It seems, however, that some new techniques should be used since there are important differences between the one- and higher-dimensional case. For example, the extension of piecewise nonparametric modeling to higher dimensions is difficult in practice. This additional step is still under study.

Acknowledgments

The authors wish to thank M. Simioni for very helpful discussions and suggestions.

References

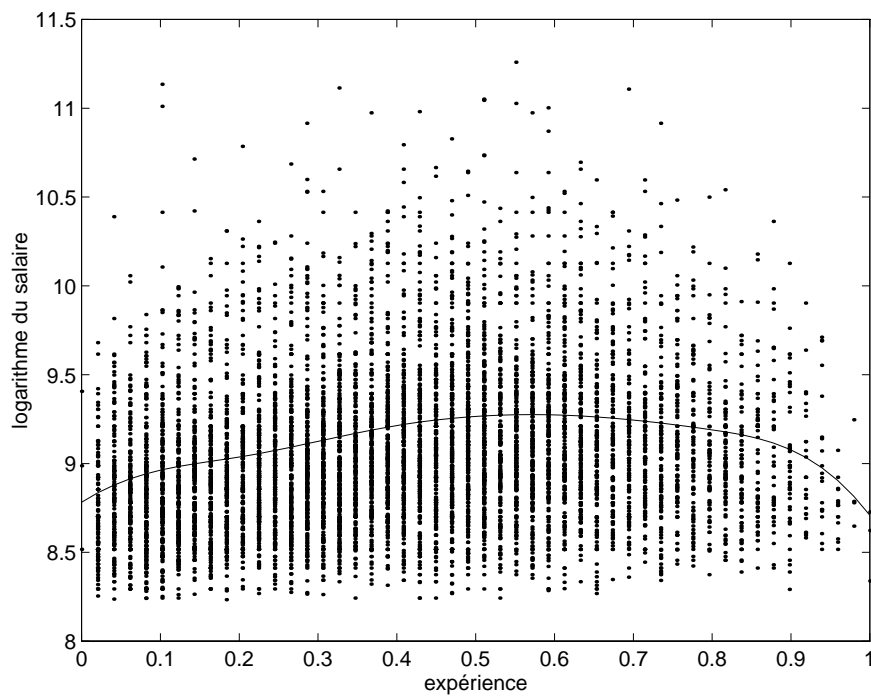
- [1] G.G. Agarwal and W.J. Studden. Asymptotic integrated mean square error using least squares and bias minimizing splines. *The Annals of Statistics*, 8(6):1307–1325, 1980.
- [2] R.K. Beatson. Monotone and convex approximation by splines: error estimates and a curve fitting algorithm. *SIAM J. of Math. Analysis*, 19(4), 1982.
- [3] R.L. Berger. Uniformly more powerful tests for hypotheses concerning linear inequalities and normal means. *Journal of The American Statistical Association*, 84(405):192–199, 1989.
- [4] E.R. Berndt. *The practice of econometrics*. Addison-Wesley Publishing Company, 1991.
- [5] P.G. Ciarlet and J.M. Thomas. Exercices d’analyse matricielle et d’optimisation. *Collection mathématiques appliquées pour la maîtrise*. Masson, 1982.
- [6] C.A.T. Diack. Testing convexity. *Compstat’96: Proceedings of Physica Verlag*, 1996.
- [7] H. Dierckx. An algorithm for cubic spline fitting with convexity constraints. *Computing*, 24:349–371, 1980.
- [8] W. Hardle and E. Mammen. Comparing nonparametric versus parametric regression fits. *The Annals of Statistics*, 21(4):1926–1947, 1993.
- [9] C. Kelly and J. Rice. Monotone smoothing with application to dose-response curves and the assessment of synergism. *Biometrics*, 46:1071–1085, 1990.
- [10] W. Schlee. Nonparametric test of the monotony and convexity of regression. *Nonparametric Statistical Inference*, 2:823–836, 1980.
- [11] L. Schumaker. *Spline function: Basic theory*. John Wiley, New York, 1981.
- [12] M. Simioni. Fonctions de gains: une approche non paramétrique. *To appear in INSEE Methods*, 1995.
- [13] A.J. Yatchew. Nonparametric regression tests based on least squares. *Econometrics Theory*, 8:435–451, 1992.

Table 4.1

Variables	Mean	Standard-deviation	Minimum	Maximum
Wage	9437.97	5458.56 0	3766.67	77583.33
Experience	19.80	11.10	0.00	49.00

Table 4.2

λ_n	$k_n = 3$	$k_n = 9$	$k_n = 15$	$k_n = 21$	$k_n = 27$	k_{n,λ_n}^{gcv}
0	1	0	0	0	0	0
$5 \cdot 10^{-3}$	0	1	1	0	0	0
$5 \cdot 10^{-3.5}$	1	1	1	1	0	1
$5 \cdot 10^{-4}$	1	1	1	1	1	1
$5 \cdot 10^{-4.5}$	1	1	1	1	0	1
$5 \cdot 10^{-5}$	1	1	0	0	0	0



$$\lambda_n = 5 \cdot 10^{-4}, k_n = 4$$

Figure 1: Hybrid cubic spline estimator.