Econometric Theory, **28**, 2012, 363–386.
doi:10.1017/S0266466611000351

SPECIFICATION TEST FOR CONDITIONAL DISTRIBUTION WITH FUNCTIONAL DATA

FREDERIC FERRATY
Université Paul Sabatier

ALEJANDRO QUINTELA-DEL-RÍO
Universidad de A Coruña

PHILIPPE VIEU
Université Paul Sabatier

In this paper we construct a statistic to test a specific form of the conditional density function. The main point of this work is to consider a functional explanatory variable, and the statistic is constructed following recent advances in nonparametric functional data analysis. The asymptotic behavior of the test statistic is studied under both the null hypothesis and some local alternative hypothesis. Then, the finite sample behavior of the method is studied through simulated examples. This paper is one of the first in the setting of nonparametric specification tests when functional data are involved.

1. INTRODUCTION

Functional data are appearing more and more often in many fields of applied statistics such as medicine, biostatistics, environmetrics, chemometrics, bioinformatics, and econometrics. An emerging challenging question for the statistical community is to propose statistical tools for their treatment. Previous approaches in these directions were mainly considering parametric-type models/methods, in the sense that the complexity of the functional structure of the variables (which are basically objects lying in infinite dimensional space) is balanced by restrictive modeling conditions on the underlying problem. A typical well-known example is the functional linear regression model. This model is appealing in that it makes it possible to deal with functional covariates, although it needs a restrictive linearity assumption on the model, which is much more difficult to validate in the presence of functional covariates than in standard multivariate regression; see the monograph by Ramsay and Silverman (2005).

We are thankful for the contributions of the editor and two anonymous referees, who helped us to improve the previous version of the paper. The authors are also grateful to all the participants of the working group STAPH (http://www.lsp.ups-tlse.fr/staph) in Toulouse for their stimulating and helpful comments. This work was supported by the Spanish Ministry of Science and Technology, under research project MTM2006-03523. Address correspondence to Alejandro Quintela-del-Río, Universidad de A Coruña, Campus de Elviña s/n, 15071 A. Coruña, Spain; e-mail: aquintela@udc.es.

© Cambridge University Press 2011

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

In recent years a new methodology, based on nonparametric ideas, has emerged for functional data. Many statistical problems involving functional variables can be attacked nonparametrically. The monograph by Ferraty and Vieu (2006) discusses the state of the art for this nonparametric functional data analysis framework. Such nonparametric functional data approaches cover several estimation problems (including regression, conditional density, quantile, mode, etc.) as well as curves discrimination or curves classification. Except for the recent works by Delsol (2008) on regression, there exist no contributions in the testing framework. Furthermore, nonparametric techniques have been widely used for many different testing purposes in a standard finite dimensional statistical setting. Among the numerous publications in this field, note the works of Delgado and González-Manteiga (2001), Guerre and Lavergne (2002), Fan and Li (1996), or Härdle and Mammen (1993). For a discussion on the use of smoothing techniques in testing procedures, see the monograph by Hart (1997). As discussed before, this has not yet been studied in the infinite dimensional framework, and the aim of this paper is to propose a procedure for testing some special parametric form for the conditional density with functional explanatory variable.

Section 2 briefly presents the wide scope of fields (with special emphasis on econometrics) for which functional data modeling and conditional density estimation are important. Section 3 defines the nonparametric estimate and states the testing problem. Section 4 presents the test statistic, which is constructed and studied theoretically under the null and a local alternative hypotheses. Section 4 includes theoretical examples appearing in econometrics (diffusion processes, fractional Brownian motion, etc.) and highlights the importance of the topological structure endowing the functional space of the variables. Section 5 shows the feasibility of the procedure through finite simulated samples. Section 6 discusses a few open problems. Technical details appear in the Appendix.

2. ON FUNCTIONAL DATA MODELING

2.1. Modeling Functional Data Sets

Suppose that a real variable X is observed on a grid

$$X(z_1), \dots X(z_p), \quad a = z_1 \leq z_2 \dots \leq z_p = b. \tag{1}$$

When p is large, the high dimensionality of the problem and the strong correlation expected between successive observations make the standard tools not behave correctly, and these kinds of data cannot be treated efficiently using standard multivariate statistical tools. Functional data modeling consists in looking at the data in a functional way, that is, each datum is a continuous curve

$$\chi = \{X(z), a \leq z \leq b\}, \tag{2}$$

and functional statistics aim to develop specific methods to deal with samples $\chi_1, \dots \chi_n$ of such data. From a mathematical perspective, it is assumed that the χ_i

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

are independent random variables taking values in some abstract space of infinite dimension. Basically, the main idea behind functional data analysis is to say that each datum as in (2) states a statistical problem based on a single (functional) object, while a standard multivariate look (1) provides a (very) high statistical problem. Section 4.3 will discuss the role of the topological structure endowing this abstract space and will show how small ball probability considerations could serve to theoretically validate this dimensionality reduction idea.

2.2. Functional Data and Econometrics

Given the progress achieved in collecting and storing data, most applied scientists now have to deal with variables observed in finer and finer grids, and the functional modeling each day becomes more and more important in practice. From a historical perspective, chemometrics is one of the first fields where functional data modeling has been used. Because direct chemical analysis of any substance is always very expensive and time consuming, people prefer to look at the spectrometric curve observed for this substance. The spectrometric analysis gives the percentage of light absorbance (here denoted as X) as a function of the wavelength (here denoted as z) near the infrared spectrum. The precision of the spectrometric tools now available on the market provides very fine grids (p can be several hundred) and so one has really observed a continuous curve (let us say χ) as described in (2). Many applications in this direction, particularly in food quality controls, have been carried out by functional data modeling (see Ferraty and Vieu, 2006, for more details and references). As discussed by Ramsay and Silverman (2002), many other science fields now deal with this type of data (climatology, environmetrics, biostatistics, etc.).

Of course econometrics is not an exception to this rule, and the high potential of interactions between econometrics and functional data modeling is highlighted by the following selected set of recent works. In Bugni, Hall, Horowitz, and Neumann (2009), functional data ideas are used for modeling unemployment durations. Jank and Shmuel (2006) use functional modeling for the analysis of electronic transactions, and Laukaitis (2008) for modeling and predicting cash flow and intensity of transactions in credit card payment systems. Reddy and Dass (2006) consider functional prices, while Ramsay and Ramsay (2002) study index of production in dynamics analyses. Ramsay, Hooker, and Graves (2009) provide functional statistical analysis for input-output data from the oil industry. Functional data ideas are also useful for modeling and analyzing volatility (Alva, Romo, and Ruiz, 2009). The aim of our work is to provide new insights on the conditional density function when the data are curves.

2.3. Functional Data and Conditional Density

In many problems, a variable χ is associated with some response variable Y. It is well known that conditional density is an interesting tool for depicting the

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

links between χ and Y, and may provide (via conditional mode) good alternative tools to the standard conditional mean regression method for predicting Y from χ (see Linton and Sancetta (2009), where they used kernel smoothing to estimate conditional distributions given an infinite past). Many theoretical studies have been performed and various statistical tools have been proposed when the explanatory variable is not functional. Once again, econometrics is a field for which conditional density is of great importance (see Chen, Linton, and Robinson (2000) for theoretical advances, Cai and Wang (2008) for an application of conditional density estimation in value-at-risk analysis, Hansen (1994) for a financial application, and Racine (2008) or Fan (2005) for general bibliographic sets of references). These works have been carried out for standard multivariate explanatory variables, and the aim of our work is to provide advances in this direction when the explanatory variable χ is a curve. From a predictive perspective, in the special case when χ represents the past of some real valued process, the functional framework makes it possible to take into account the whole past of the process (on some path of length $[a, b]$) while the multivariate approach obliges, because of a dimensionality problem, to reduce the information to a very small number of past values. This has direct practical impact in econometrics (for instance, in price analysis, risk analysis, volatility analysis, and value-at-risk analysis).

3. THE FUNCTIONAL CONDITIONAL DENSITY CONTEXT

3.1. The Conditional Density Nonlinear Operator

Let $\{(\mathcal{X}_i, Y_i), i = 1, \ldots n\}$ be a sample of n independent random pairs, each being distributed as $(\mathcal{X}, Y)$. This paper deals with the conditional density function of Y given $\mathcal{X}$, in a situation where the explanatory variable $\mathcal{X}$ is allowed to be of a functional nature. More precisely, we consider that $\mathcal{X}$ is a random variable valued in some semimetric space (E, d) that is (possibly) of infinite dimension. The response variable Y is scalar (i.e., Y takes its values in $\mathbb{R}$). The conditional cumulative distribution of Y given $\mathcal{X}$ is defined for any $y \in \mathbb{R}$ and any $\chi \in E$ by

$$F(y/\chi) = P(Y \leq y/\mathcal{X} = \chi),$$

while the conditional density, denoted by $f(y/\chi)$, is defined to be the density of this distribution with respect to Lebesgue measure on $\mathbb{R}$. In all of the following, we are implicitly assuming that both F and f exist and are well defined (see Ferraty and Vieu, 2006, Ch. 5.2, for more on this point).

We wish to estimate it from a statistical sample $\{(\mathcal{X}_i, Y_i), i = 1, \ldots n\}$. At this stage, it is worth noting that in the functional setting (that is, when E is of infinite dimension), f and F are not standard mathematical objects. Because they are defined on infinite dimensional spaces, we will use the terminology *operators* (rather than functions). Moreover, because we wish to address the statistical problems in a free-parameter way, it is out of purpose to make any linearity assumption, and to emphasize this point we will use the terminology *nonlinear operators* for f and

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

F. In addition, smoothness assumptions will only be made on these nonlinear operators (see Condition H.1 below), leading us to a nonparametric setting.

3.2. The Functional Kernel Conditional Density Estimate

Following Ferraty, Laksaci, and Vieu (2006), the conditional density operator $f(y/\chi)$ can be estimated by means of kernel smoothing ideas. Such an estimate can be defined as

$$\hat{f}_n(y/\chi) = \frac{\frac{1}{g}\sum_{i=1}^{n} K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right) K_0\left(\frac{y-Y_i}{g}\right)}{\sum_{i=1}^{n} K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)}, \tag{3}$$

where K and K_0 are kernel functions, and where g and h are sequences of smoothing parameters (see Conditions H.2 and H.3 below). Asymptotic properties (consistency with rates) as well as practical interest of this estimate for prediction problems by means of conditional mode have been studied in Ferraty et al. (2006). As in any nonparametric problem involving functional data, the behavior of the estimate is controlled by the concentration properties of the functional variable $\mathcal{X}$. In other words, the small ball probability function

$$\varphi_\chi(\epsilon) = P(\mathcal{X} \in \mathcal{B}(\chi,\epsilon)) \tag{4}$$

will enter directly into the expression of the rates of convergence of the estimate. Here $\mathcal{B}(\chi,\epsilon) = \{\chi' \in E, d(\chi,\chi') \leq \epsilon\}$ is the ball related with the topology associated to the semimetric d. See Section 4.3 for more on this point.

From now on, let us consider that χ is a fixed element of the functional space E, and let us give some technical hypotheses that will be necessary later on. The nonparametric model for the conditional density operator is defined with some assumptions.

Condition H.1a. The function $f(y/\cdot)$ is Lipschitz continuous in a neighborhood of χ.

Condition H.1b. $f(\cdot/\chi)$ has four bounded continuous derivatives on $\mathbb{R}$.

Concerning the kernel functions involved in the density estimates, we will assume

Condition H.2a. $K_0(\cdot)$ is a symmetric density, $\int K_0^2(v)dv < \infty$ and $0 < \int v^4 K_0(v)dv < \infty$.

Condition H.2b. K is the asymmetrical box kernel (i.e., $K(u) = I_{[0,1]}(u)$).

Extension to other kernels would be possible (Ferraty and Vieu, 2006, Ch. 4), but would lead to tedious additional technical proofs that would mask the main goal of this paper. Finally, the bandwidths should verify, together with the small ball probability function φ_χ, the following conditions.

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

Condition H.3. $g \to 0,\ h = o(g^2),\ n\varphi_\chi(h)^3 \to \infty,\ n\varphi_\chi(h)g^8 \to 0,\ n\varphi_\chi(h)g^6 \to \infty$.

3.3. The Testing Problem

Any information on the conditional density function is always of great practical interest, and the statistical literature in this field is quite extensive when the explanatory variable is real or multivariate. In particular, the question of testing some special feature of the conditional density has been widely investigated, and one possible way of constructing test statistics is to look at how a nonparametric estimate is far (or not) from a given fixed family of density functions. Just to stay with a few recent references, we would like to cite the works by Hyndman and Yao (2002) and Fernandes and Grammig (2005), in which some testing procedures are proposed using nonparametric estimates in a finite dimensional framework. Naturally, the same problem appears when one is working with functional data, and the aim of this paper is to use the kernel conditional density estimate $\hat{f}_n$ to construct statistical procedures to test whether the true conditional density is equal to some specific one; or, more generally, if it belongs to some specific given parametric form.

Recall that χ is a fixed functional element of E. Section 4 proposes a statistic test for the null hypothesis

$$H_0 : \exists \theta_0,\ f(y/\chi) = f_{\theta_0}(y/\chi), \tag{5}$$

where $\{f_\theta(./\chi),\ \theta \in \Theta\}$ is a specific parametric family of densities. The normality of the statistic under H_0 is stated as long as, roughly speaking, there exists an estimate of the conditional density function under the parametric model H_0 that has better rates of convergence than the nonparametric one. This is a quite unrestricted condition, and this is why the space Θ of the parameters of the null model is deliberately unspecified, and we just suppose that we have at hand some parametric estimate $f_{\hat{\theta}_0}$ that satisfies

Condition H.4. There exists a sequence $U_n = O\left(\left(n\varphi_\chi(h)\right)^{-1/2}\right)$, such that under H_0,

$$\forall y \epsilon S_w,\ \left| f_{\theta_0}(y/\chi) - f_{\hat{\theta}_0}(y/\chi) \right| = o(U_n), \qquad \text{a.s.}$$

Section 4 will also state the divergence under a local alternative of the form

$$H_1 : \inf_{\theta \in \Theta} |f(y/\chi) - f_\theta(y/\chi)| = \varepsilon_n s(y), \tag{6}$$

under the following general restriction.

Condition H.5. $\sup_{y \in S_w} s(y) < \tau < \infty,\ \int f''(y)s(y)w(y)dy \neq 0$, and $\varepsilon_n \sqrt{n\varphi_\chi(h)} \to \infty$.

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

4. THE TESTING PROCEDURE

4.1. Construction of the Statistic

In this section we will test H_0 against H_1 as defined before. Here, $w(\cdot)$ is some known bounded weight function with compact support. The idea is based on the comparison between the nonparametric kernel functional estimate $\hat{f}_n(\cdot/\chi)$ and the parametric estimate $f_{\hat{\theta}_0}$ adapted to the hypothesis H_0. Consider the quantity

$$I_n = \int (\hat{f}_n(y/\chi) - f_{\hat{\theta}_0}(y/\chi))^2 w(y)dy, \tag{7}$$

where $w(\cdot)$ is a known bounded weight function with compact support S_w. We have that

$$I_n = \int \left[\frac{\frac{1}{g}\sum_{i=1}^n K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right) K_0\left(\frac{y-Y_i}{g}\right)}{\sum_{i=1}^n K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)} - f_{\hat{\theta}_0}(y/\chi) \right]^2 w(y)dy.$$

Rewriting I_n as

$$I_n = \int \left[\frac{\sum_{i=1}^n K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right) \frac{1}{g} K_0\left(\frac{y-Y_i}{g}\right)}{\sum_{i=1}^n K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)} - f_{\hat{\theta}_0}(y/\chi) \frac{\sum_{i=1}^n K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)}{\sum_{i=1}^n K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)} \right]^2 w(y)\,dy$$

$$= \int \left[\frac{\frac{1}{n\varphi_\chi(h)}\sum_{i=1}^n K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right) \left(\frac{1}{g} K_0\left(\frac{y-Y_i}{g}\right) - f_{\hat{\theta}_0}(y/\chi)\right)}{\frac{1}{n\varphi_\chi(h)}\sum_{i=1}^n K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)} \right]^2 w(y)\,dy,$$

the nullity of I_n can be equivalent to the nullity of the quantity

$$I'_n = \int \left[\frac{1}{n\varphi_\chi(h)} \sum_{i=1}^n K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right) \left(\frac{1}{g} K_0\left(\frac{y-Y_i}{g}\right) - f_{\hat{\theta}_0}(y/\chi)\right) \right]^2 w(y)\,dy, \tag{8}$$

Finally, our testing procedure is based on the following normalized version of the statistic

$$J_n = \frac{\sqrt{\frac{n\varphi_\chi(h)}{g^4}}\left(I'_n - g^4 B\right)}{\sqrt{V}}, \tag{9}$$

where

$$B = \frac{1}{4}\left(\int v^2 K_0(v)dv\right)^2 \left(\int f''(y/\chi)^2 w(y)\,dy\right)$$

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

and

$$V = \left(\int f(y/\chi) f''(y/\chi)^2 w(y)^2 dy \right.$$
$$\left. - \left(\int f(y/\chi) f''(y/\chi) w(y) dy \right)^2 \right) \left(\int v^2 K_0(v) dv \right)^2.$$

Note that V can also be written as

$$V = \mathrm{Var}(f''(Y/\chi) w(Y) | \mathcal{X} = \chi) \left(\int v^2 K_0(v) dv \right)^2.$$

Throughout the remaining part of the paper it is, of course, implicitly assumed that $V > 0$.

4.2. Some Asymptotics

The asymptotic distribution of the statistic J_n is stated in the next theorem.

THEOREM 1. *If the Conditions H.1–H.4 hold, then under H_0 we have*

$$J_n \longrightarrow N(0,1), \quad as\ n \to \infty.$$

The divergence property of the testing procedure under the local alternative introduced in Condition H.5 is now given.

THEOREM 2. *If the assumptions H.1–H.5 hold, then under H_1 we have*

$$|J_n| \longrightarrow \infty, \quad in\ probability.$$

The proofs of Theorems 1 and 2 are reported in the Appendix.

4.3. On the Semimetric and the Small Ball Probability

The small ball probability function φ_χ plays a major role in nonparametric estimation procedures (Ferraty and Vieu, 2006, Ch. 13). In the testing setting investigated here, such a small ball probability function appears directly in the construction of the test statistic J_n (as a normalizing factor). Such a consideration is also standard in probability theory for Gaussian continuous time processes, where the lack of any measure of reference (as could be the Lebesgue measure in finite dimensional space) forces one to characterize processes by means of their small ball probability function (rather than with some density function as in the finite dimensional case). For more, see the monograph by Bogachev (1999) and the survey paper by Li and Shao (2001). The aim of this section is to illustrate, through standard examples of Gaussian processes with great impact in econometrics, the effects of the small ball probability function on our test procedure. In addition, because the small ball probability function is directly linked with the topological

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

structure on the functional space E (that is, with the semimetric d endowing the space E), we will discuss the impact of this topological parameter d on our procedure. In the remainder of this section a, b, c, C, or C_χ are strictly positive generic real constants.

Example 1: Diffusion processes
One among the most known diffusion processes is the Ornstein-Uhlenbeck χ^{OU} one, constructed from the usual Wiener process w as the solution of the differential equation

$$d\chi^{OU}(t) = -\sigma\chi^{OU}(t)dt + dw(t), \quad \sigma > 0, \quad 0 < t \leq 1, \quad \text{and } \chi^{OU}(0) = 0.$$

For the supremum norm

$$\|x\|_{\sup} = \sup_{0\leq t\leq 1}|x(t)|,$$

the small ball probability function of the Ornstein-Uhlenbeck diffusion process has the asymptotic form

$$\varphi_\chi^{OU}(\epsilon) = P(\chi^{OU} \in \mathcal{B}_{\sup}(\chi,\epsilon)) = P(||(\chi^{OU} - \chi||_{\sup} \leq \epsilon)$$
$$\sim C_\chi e^{-\frac{C_\sigma}{\epsilon^2}}, \quad \text{as } \epsilon \sim 0,$$

where $\mathcal{B}_{\sup}$ denotes the ball for the topology associated with the supremum norm $\|.\|_{\sup}$ and where χ is a fixed deterministic element in $\mathcal{C}([0,1])$ (see Bogachev, 1999). Moreover, one could see that such an asymptotic expansion is not only true for the standard Ornstein-Uhlenbeck χ^{OU} process. Pending to general assumptions insuring that some diffusion process χ^{DIFF} has an absolutely continuous distribution with respect to the Wiener measure, one can get expressions of the type

$$\varphi_\chi^{DIFF}(\epsilon) \sim C_\chi e^{-\frac{c}{\epsilon^2}}, \quad \text{as } \epsilon \sim 0.$$

Example 2: Fractional Brownian motion
The fractional Brownian motion, which we will denote by χ^{FBM}, is also known to have the same kind of asymptotic small ball probability function. Specifically, one has that ($\delta \in]0,2[$ being the order of the fractional Brownian motion)

$$\varphi_\chi^{BFM}(\epsilon) \sim C_\chi e^{-\epsilon^{-\frac{2}{\delta}}}, \quad \text{as } \epsilon \sim 0,$$

(see Li and Shao, 2001; Thms. 3.1 and 4.6).

Example 3: The general family of exponential processes
Many time-continuous processes have the same exponential-type asymptotic small ball probability function. A process χ^{PROC} is called an exponential-type process, with respect to some abstract topology, if

$$\varphi_\chi^{PROC}(\epsilon) \sim C_\chi e^{-\frac{\log(\frac{1}{\epsilon})^a}{\epsilon^b}}, \quad \text{as } \epsilon \sim 0.$$

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

As seen before, the fractional Brownian motion as well as many diffusion processes are of exponential type when the topology is associated with the supremum norm. This is true for many other processes (possibly being of more than 1 dimension as Levy processes, fractional Brownian sheets, fractional Ornstein-Uhlenbeck process, etc.) and also for various other normed topologies (e.g., Hölder norms, f-norms). The probabilistic literature in this field is increasing each day, and currently examples can be found in Bogachev (1999) or in Li and Shao (2001).

What Is the Impact on the Functional Testing Procedure? Assume that χ^{PROC} is an exponential-type process. The form itself of the small ball probability function leads to the conclusion that the only way to satisfy Condition H.3 is to choose a very large bandwidth of the form $h \sim C(\log n)^{-b}$. This would lead to normalizing factor $\sqrt{n\varphi_\chi(h)}$ in our statistic test of the form $(\log n)^b$.

Links with Semimetric Choice. The slow rate of convergence stated just before is not surprising because we are doing infinite dimensional nonparametric statistics, and this is stressing an interesting point. Going back to the probabilistic literature discussed before, exponential-type processes are only obtained for topological structures associated with a norm. One way to overcome this problem is to consider other kinds of topologies, and this is exactly the aim of the approach developed here when we use abstract semimetric space. To further show the advantages of this semimetric approach, one can look at the following specific example. Suppose that the space E is endowed with some inner product $<;>$ with a countable basis $(e_1, \ldots, e_n, \ldots)$, and define the semimetric d on E just to be the euclidean metric between the k (k being fixed for simplicity here) first projections on this basis:

$$\forall(\chi_1, \chi_2) \in E^2,\ d_k(\chi_1, \chi_2) = \sqrt{\sum_{j=1}^{k} < \chi_1 - \chi_2, e_j >^2}.$$

It is easy to see that, with the new topology defined by d_k, the small ball probability is of polynomial form,

$$\varphi_\chi(\epsilon) \sim C\epsilon^k, \quad \text{as } \epsilon \sim 0.$$

This is true for any space E. So, even the previous Gaussian and diffusion processes, which have very quickly (exponentially) decreased small ball probability functions in terms of the usual normed topology, can be shown by this simple topological change to have much more tractable polynomial small ball probability. This is of great practical impact, because that means that with such a semimetric choice, one could choose the bandwidth h of the standard form $h \sim Cn^{-c}$ and arrive directly to the normalizing factor $\sqrt{n\varphi_\chi(h)}$, which has a faster rate of the form n^{-c}.

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

As a Matter of Conclusion. Finally, the discussion before highlights the fact that the semi-metric modeling, which could appear as a purely cosmetic artifact of mathematicians, is in fact a key point of the procedure. A choice of semimetric (see just before) ensuring good rates of convergence of the procedure is always possible for any kind of processes. In some sense, the semimetric modeling is a good answer to the dimensionality problems described in Section 2.1. We stop this discussion here and refer to Ferraty and Vieu (2006) for more on this point, in particular on choosing the basis on which the functional data should be projected, and also for alternative semimetric constructions (since this projection idea is just one example of a semimetric that one could use in practice).

5. SOME SIMULATIONS

5.1. Presentation of the Experiment

The Simulated Data. We simulate the process $Z(t)^* = \frac{Z(t)+Z(t+1)+Z(t+19)}{20}$, $t = 1, \ldots, N$, where $Z(t)$ are generated as $Z(t) = 0.9Z(t-1) + a(t)$, with $a(t) \sim \mathcal{N}(0, 0.1)$. We choose $N = 100 \times n$, and then we split into n paths of length 100 $\chi_i = \{\chi_i(t) = Z^*(t + 10 \times (i-1)), t = 1, \ldots, 100\}$, $i = 1, \ldots, n$. These n paths χ_i are the functional explanatory variable in our experiment. Then, the responses are generated by

$$Y_i = r(\mathcal{X}_i) + \varepsilon_i, \qquad i = 1, \ldots n, \quad \text{with} \quad r(\mathcal{X}) = \int \mathcal{X}(t)\cos(t)dt. \tag{10}$$

The null hypothesis consists in simulating the innovation errors as

$$H_0 : \quad \varepsilon \sim \mathcal{N}(0, \sigma_0).$$

Note that this hypothesis could be expressed in terms of the conditional density of Y given $\mathcal{X}$, since under H_0 this conditional distribution is just $\mathcal{N}(\mu_\chi, \sigma_0)$. The alternative H_1 is constructed by simulating the errors as a mixture of $\mathcal{N}(-d, \sigma_1)$ and $\mathcal{N}(d, \sigma_1)$ distributions.

The Estimates. The various parameters of the methods were chosen in the following ways. The kernel function K was taken to be the uniform density on $(0, 1)$, and the corresponding bandwidth h was selected by cross-validation. Then, the function K_0 was taken to be the usual Epanechnikov kernel and, to save time, the corresponding bandwidth g was fixed as $g = (\max(Y_i) - \min(Y_i))/10$. Finally, the semimetric d was taken to be the usual L_2 distances between curves. Motivations for these choices of the parameters are given in Ferraty and Vieu (2006) and in its companion R package, from which all the routines used here have been extracted. The pilot nonparametric conditional estimate $\hat{f}_n$ is now built. The parametric estimate $f_{\hat{\theta}_0}$ was taken as a normal density with, to speed up the algorithm, parameters estimated in such a way that the modes of both nonparametric and parametric estimates are the same.

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

The Experiments. The statistic test J_n was computed for various values of the parameters n, d, σ_0, and σ_1, making it possible to see the behavior of the testing procedures when both hypotheses H_0 and H_1 are changing. In each case, 100 replications of the experiment were performed. The constants appearing in the statistic J_n (corresponding to the mean and variance of I_n under H_0) were approximated by the empirical mean and variance of I_n, obtained by performing (in each situation) 500 independent preliminary replications of the experiment under H_0. An alternative procedure for that would be to use bootstrapping ideas, but the asymptotic validity of bootstrap in functional testing problems is still not proven (see, however, Ferraty, Van Keilegom, and Vieu (2010) for first advances in bootstrapping for nonparametric functional regression). In all the studies, the level of the test was fixed to be 0.95.

5.2. First Results

We start by presenting with details the specific situation when $n = 100$, when σ_0 is chosen to correspond to a signal-to-noise ratio in the null model of 0.1, and when $d = 0.3$ and $\sigma_1 = \sigma_0$. To fix the ideas, Figures 1 and 2 present, under H_0 and H_1 respectively, the nonparametric conditional density pilot estimate and the parametric Gaussian estimate.

Figure 1 (resp. Figure 2) shows that, under the Gaussian hypothesis H_0 (resp. under the alternative H_1), both estimates are closed (resp. farther) leading one to expect the statistic J_n to be small (resp. high). This is confirmed in Table 1, which gives, for 100 replications of this experiment, the percentage of acceptance of the

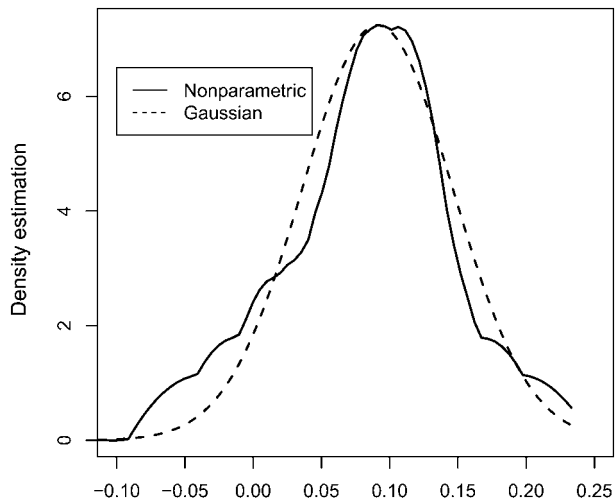


FIGURE 1. Conditional kernel and parametric estimates under H_0.

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

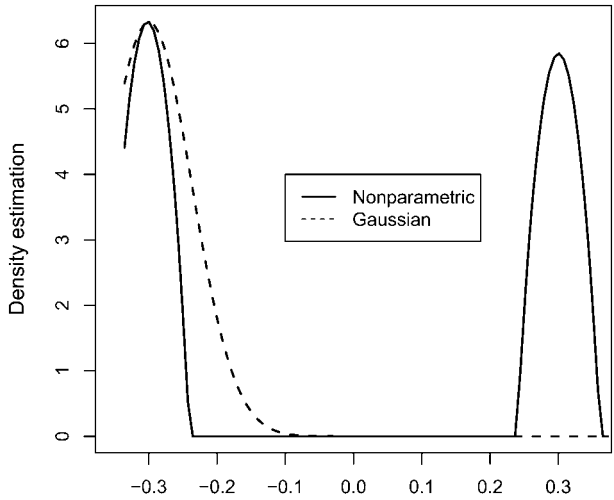


FIGURE 2. Conditional kernel and parametric estimates under H_1.

TABLE 1. Percentages of acceptance of normality assumption.

%	Reject	Accept
Under H_0	0.09	0.91
Under H_1	0.87	0.13

normal hypothesis under H_0 (first line) and under the alternative H_1 (second line). This simple pedagogical situation will be completed in the next section.

5.3. More Insight under H_0

In this section all the studies are performed under the null hypothesis H_0. We used various sample sizes, and we tried various null hypotheses by changing the value of σ_0. Precisely, denoting the signal-to-noise ratio by snr we used values of σ_0 of the form $\sigma_0 = c \times 0.1 \times snr$, and we express the results as a function of c. To fix ideas, note that the case $c = 1$ corresponds to the situation depicted in Section 5.2. Tables 2 and 3 show that, even for moderate sample sizes and for high signal-to-noise ratio, the testing procedure is good, since the percentage of acceptance is always closed from the true theoretical level.

5.4. More Insight under H_1

In this section the studies are performed under the alternative hypothesis H_1. We look at how the power of the test procedure changes. We choose various null

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

TABLE 2. Results under H_0 when σ_0 varies (with $n = 100$).

c	1	2	3	4	5
% acceptance of H_0	0.91	0.90	0.95	0.98	1

TABLE 3. Results under H_0 when the sample size n varies (with $c = 1$).

Sample size	% acceptance of H_0	% acceptance of H_1
100	0.91	0.13
200	0.93	0.03
500	0.98	0

TABLE 4. Results under H_1 when d varies ($n = 100$ and $c = 1$).

d	0.1	0.25	0.3	0.5
% Rejection of H_0	1	0.96	0.94	0.38

TABLE 5. Results under H_1 when σ_1 varies ($n = 100$ and $d = 0.1$).

c	1	2	4	5	10	20
% Rejection of H_0	1.0	0.93	0.91	0.86	0.91	0.86

hypotheses H_1, constructed from different values of d and σ_1. The sample size is fixed to $n = 100$. As before, the values of σ_1 are chosen of the form $\sigma_1 = c \times 0.1 \times snr$. Tables 4 and 5 show the evolution of the percentage of rejection as a function of d and σ_1, respectively.

For moderate sample size and high signal-to-noise ratio, the power of the test is rather high. The only counterexample is when the distance between both modes of the density is high (see Table 4 and large value of d) because the nonparametric estimate tends to oversmooth the distribution. Except for this situation, which is rather unusual in practice (and easy to detect empirically), the procedure has good power.

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

6. CONCLUSIONS AND TRACKS FOR FUTURE RESEARCH

This paper shows that by using a nonparametric kernel-type conditional density estimate, one could construct a statistic to test a specific form of the true conditional density estimate. Among other possible uses, this result may be useful in regression problems for testing normality of the errors. The major point of this paper is to allow the explanatory variable to take values in some infinite dimensional space (opening the way for many applications involving curves). It turns out that this functional feature appears through the small ball probability function of the functional variable. As far as we know, this is one of the first papers in which nonparametric estimates are used for model specification testing with functional variables. We would like to stress the fact that the methodology that is developed here for conditional density could be followed (at least in its main lines) in many other testing problems involving functional variables (such as regression, time series prediction, conditional cumulative distribution, etc.). Another interesting problem would be, as discussed in Section 5, the study of bootstrap procedures to avoid the tedious direct estimation of the constant terms involved in the asymptotic distribution of the statistic. Also, as pointed out by an editor, an appealing challenge would be to consider a uniform (on χ) null hypothesis. This would need to construct integrated (over the functional variable) statistics and would require more technical insights. For the specific problem of regression, some advances (both on bootstrap and on integrated statistics) are discussed in the recent work by Delsol (2008) and could serve as a basis for such extensions in the conditional density setting.

Undoubtedly, the development of such nonparametric testing procedures for functional data is a real challenge for the statistical community, and we hope that our paper will encourage others to perform more research in this direction.

REFERENCES

Alva, K., J. Romo, & E. Ruiz (2009) Modeling Intra-daily Volatility by Functional Data Analysis: An Empirical Application to the Spanish Stock Market. Working Paper 09-28, Statistics and Econometrics Series, University Carlos III.

Bogachev, V.I. (1999) *Gaussian Measures.* Mathematical surveys and monographs 62. American Mathematical Society.

Bugni, F.A., P. Hall, J.L. Horowitz, & G.R. Neumann (2009) Goodness-of-fit tests for functional data. *Econometrics Journal* 12, 1–18.

Cai, Z. & X. Wang (2008) Nonparametric estimation of conditional VaR and expected shortfall. *Journal of Econometrics* 147, 120–130.

Chen, X., O. Linton, & P. Robinson (2000) The estimation of conditional densities. In M.L. Puri (ed.), *Asymptotics in Statistics and Probability: Papers in Honor of George Roussas*, pp. 71–84. VSP International Science Publishers.

De Jong, P. (1987) A central limit theorem for generalized quadratic forms. *Probability Theory and Related Fields* 75**,** 261–277.

Delgado, M. & W. González-Manteiga (2001) Significance testing in nonparametric regression based on the bootstrap. *Annals of Statistics* 29, 1469–1507.

Delsol, L. (2008) *Regression sur variables fonctionnelles: Estimation, tests de structure et estimation*. Ph.D. thesis, Universite Paul Sabatier.

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

Fan, J. (2005) A selective overview of nonparametric methods in financial econometrics. *Statistical Science* 20, 317–337.
Fan, Y. & Q. Li (1996) Consistent model specification tests: Omitted variables and semi-parametric functional forms. *Econometrica* 64, 865–890.
Fernandes, M. & J. Grammig (2005) Nonparametric specification tests for conditional duration models. *Journal of Econometrics* 127, 35–68.
Ferraty, F., A. Laksaci, & P. Vieu (2006) Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statistical Inference for Stochastic Processes* 9, 47–76.
Ferraty, F., I. Van Keilegom, & P. Vieu (2010) On the validity of the bootstrap in nonparametric functional regression. *Scandinavian Journal of Statistics* 37, 286–306.
Ferraty, F. & P. Vieu (2006) *Nonparametric Functional Data Analysis*. Springer Series in Statistics. Springer.
Guerre, E. & P. Lavergne (2002) Optimal minimax rates for nonparametric specification testing in regression models. *Econometric Theory* 18, 1139–1171.
Hansen, B. (1994) Autoregressive conditional density estimation. *International Economic Review* 35, 705–730.
Härdle, W. & E. Mammen (1993) Comparing parametric versus nonparametric regression fits. *Annals of Statistics* 21, 1926–1947.
Hart, J. (1997) *Nonparametric Smoothing and Lack of Fit Tests.* Springer.
Hyndman, R.J. & Q. Yao (2002) Nonparametric estimation and symmetry tests for conditional density functions. *Journal of Nonparametric Statistics* 14, 259–278.
Jank, W. & G. Shmuel (2006) Functional data analysis in electronic commerce research. *Statistical Science* 21, 155–166.
Laukaitis, A. (2008) Functional data analysis for cash flow and transactions intensity continuous-time prediction using Hilbert-valued autoregressive processes. *European Journal of Operational Research* 185, 1607–1614.
Li, W.V. & Q.M. Shao (2001) Gaussian processes: Inequalities, small ball probabilities and applications. In C.R. Rao and D. Shanbhag (eds.), S*tochastic Processes: Theory and Methods. Handbook of Statistics* 19. North-Holland.
Linton, O. & A. Sancetta (2009) Consistent estimation of a general nonparametric regression function in time series. *Journal of Econometrics* 152, 70–78.
Racine, J.S. (2008) Nonparametric econometrics: A primer. *Foundations and Trends in Econometrics* 3, 1–88.
Ramsay, J.O., G. Hooker, & S. Graves (2009) *Functional Data Analysis with R and MATLAB*. Springer.
Ramsay, J.O. & J.B. Ramsey (2002) Functional data analysis of the dynamics of the monthly index of nondurable goods production. *Journal of Econometrics* 107, 327–344.
Ramsay, J.O. & B. Silverman (2002) *Applied Functional Data Analysis, Methods and Case Studies.* Springer Series in Statistics. Springer.
Ramsay, J.O. & B. Silverman (2005) *Functional Data Analysis (2nd ed.).* Springer Series in Statistics. Springer.
Reddy, S.K. & M. Dass (2006) Modeling on-line art auction dynamics using functional data analysis. *Statistical Science* 21, 179–193.

APPENDIX: Proofs

From now on, C and C' denote generic real positive constants. We start with two lemmas that are in a general self-contained form and that could be helpful in other contexts.

Two General Results. Note that the following two lemmas are stated without making any reference to the testing hypotheses and so they could be used throughout either under H_0 or H_1.

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

LEMMA A. *Under the Conditions H.1–H.3, we have*

$$\mathrm{E}\left[K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)\left(\frac{1}{g}K_0\left(\frac{y-Y_i}{g}\right)-f(y/\chi)\right)\right]$$
$$=\frac{1}{2}\varphi_\chi(h)\left(\int v^2K_0(v)dv\right)f''(y/\chi)g^2+O\left(\varphi_\chi(h)g^4\right).$$

LEMMA B. *Under the Conditions H.1–H.3, we have*

$$\mathrm{E}\left[K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)\left(\frac{1}{g}K_0\left(\frac{y-Y_i}{g}\right)-f(y/\chi)\right)^2\right]$$
$$=\varphi_\chi(h)\left(\int K_0^2(v)dv\right)f(y/\chi)\frac{1}{g}+O\left(\varphi_\chi(h)\right).$$

Proof of Lemma A. Put

$$\mathrm{E}\left[K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)\left(\frac{1}{g}K_0\left(\frac{y-Y_i}{g}\right)-f(y/\chi)\right)\right]$$
$$=\mathrm{E}\left[K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)\mathrm{E}\left[\left(\frac{1}{g}K_0\left(\frac{y-Y_i}{g}\right)-f(y/\chi)\right)|\mathcal{X}_i\right]\right]$$
$$=\varphi_\chi(h)\int\left(\frac{1}{g}K_0\left(\frac{y-u}{g}\right)-f(y/\chi)\right)(f(u/\chi)+O(h))du,$$

the last equality coming from Conditions H.1a and H.2b. Now, writing $(y-u)/g=v$, using a fourth-order Taylor decomposition and Condition H.3, the last term becomes equal to

$$\frac{1}{2}\varphi_\chi(h)\left(\int v^2K_0(v)dv\right)f''(y/\chi)g^2+O(\varphi_\chi(h)g^4).$$ ■

Proof of Lemma B. Applying again the properties of the conditioned expectation and Condition H.2b, we have that

$$\mathrm{E}\left[K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)\left(\frac{1}{g}K_0\left(\frac{y-Y_i}{g}\right)-f(y/\chi)\right)^2\right]$$
$$=\varphi_\chi(h)\int\left(\frac{1}{g}K_0\left(\frac{y-u}{g}\right)-f(y/\chi)\right)^2f(u/\chi)du$$
$$=\varphi_\chi(h)\left(\frac{1}{g}\int K_0^2(v)dv\right)f(y/\chi)+O\left(\varphi_\chi(h)\right).$$

The last equality is coming once again by using integration by substitution and Taylor decomposition. ■

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

Proofs of Results from Section 4. The study of the statistic J_n will be based on the decomposition

$$I_n' = T_{1,n} + T_{2,n} + T_{3,n} + 2T_{4,n} + 2T_{5,n}, \tag{A.1}$$

where

$$T_{1,n} = \frac{1}{n^2\varphi_\chi^2(h)} \int \left[\sum_{i=1}^n K^2\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right) \left(\frac{1}{g}K_0\left(\frac{y-Y_i}{g}\right) - f(y/\chi)\right)^2 \right] w(y)dy,$$

$$T_{2.n} = \frac{1}{n^2\varphi_\chi^2(h)} \int \left[\sum\sum_{i\neq j} K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right) K\left(\frac{d(\chi,\mathcal{X}_j)}{h}\right) \times \left(\frac{1}{g}K_0\left(\frac{y-Y_i}{g}\right) - f(y/\chi)\right)\left(\frac{1}{g}K_0\left(\frac{y-Y_j}{g}\right) - f(y/\chi)\right)\right] w(y)dy,$$

$$T_{3,n} = \frac{1}{n^2\varphi_\chi^2(h)} \int \left[\sum_{i=1}^n K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right) (f(y/\chi) - f_{\hat{\theta}_0}(y/\chi)) \right]^2 w(y)dy,$$

$$T_{4,n} = \frac{1}{n^2\varphi_\chi^2(h)} \int \left[\sum_{i=1}^n K^2\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right) \left(\frac{1}{g}K_0\left(\frac{y-Y_i}{g}\right) - f(y/\chi)\right) \times (f(y/\chi) - f_{\hat{\theta}_0}(y/\chi)) \right] w(y)dy$$

and

$$T_{5,n} = \frac{1}{n^2\varphi_\chi^2(h)} \int \left[\sum\sum_{i\neq j} K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right) K\left(\frac{d(\chi,\mathcal{X}_j)}{h}\right) \times \left(\frac{1}{g}K_0\left(\frac{y-Y_i}{g}\right) - f(y/\chi)\right)(f(y/\chi) - f_{\hat{\theta}_0}(y/\chi)) \right] w(y)dy.$$

Proof of Theorem 1. Theorem 1 will be a direct consequence of the two following results. The first lemma studies the limiting distribution of the leading terms in the decomposition of I_n'; the second lemma shows the negligibility of the remaining terms.

LEMMA 1. *Under the Conditions H.1–H.3, we have*

$$T_{1,n} = o_p\left(\sqrt{\frac{g^4}{n\varphi_\chi(h)}}\right) \tag{A.2}$$

and

$$\sqrt{\frac{n\varphi_\chi(h)}{g^4}}\left(\frac{T_{2,n} - g^4 B}{\sqrt{V}}\right) \longrightarrow N(0,1). \tag{A.3}$$

LEMMA 2. *Under the conditions of Theorem 1, we have*

$$T_{3,n}, T_{4,n}, T_{5,n} = o_p\left(\sqrt{\frac{g^4}{n\varphi_\chi(h)}}\right).$$

These two lemmas will be proven later in this Appendix. ■

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

Proof of Theorem 2. Note that Lemma 1 is valid independently of the null hypothesis. Therefore, Theorem 2 follows directly from the definitions (9) and (A.1), from Lemma 1, and from the following result.

LEMMA C. *Under the conditions of Theorem 2, the following results hold:*

$$\sqrt{\frac{g^4}{n\varphi_\chi(h)}} = o(|\mathrm{E}T_{5,n}|), \tag{A.4}$$

$$\mathrm{Var}(T_{5,n}) = o\left((\mathrm{E}T_{5,n})^2\right), \tag{A.5}$$

$$T_{3,n}, T_{4,n} = o_p\left(|\mathrm{E}T_{5,n}|\right). \tag{A.6}$$

Lemma 3 will be proven later in this Appendix. ■

Proofs of Technical Lemmas.

Proof of Lemma 1. It is sufficient to prove the four following results.

$$\mathrm{E}\left(T_{1,n}\right) = o\left(\sqrt{\frac{g^4}{n\varphi_\chi(h)}}\right), \tag{A.7}$$

$$\mathrm{Var}(T_{1,n}) = o\left(\frac{g^4}{n\varphi_\chi(h)}\right), \tag{A.8}$$

$$\sqrt{\frac{n\varphi_\chi(h)}{g^4}}\left(T_{2,n} - \mathrm{E}\left(T_{2,n}\right)\right) \longrightarrow N(0, \sqrt{V}), \tag{A.9}$$

and

$$\mathrm{E}\left(T_{2,n}\right) = Bg^4 + o\left(\sqrt{\frac{g^4}{n\varphi_\chi(h)}}\right). \tag{A.10}$$

Proof of (A.7). Using Lemma B, we have that

$$\mathrm{E}(T_{1,n}) = C\frac{1}{n^2\varphi_\chi^2(h)} n\varphi_\chi(h)\frac{1}{g} + O\left(\frac{1}{n\varphi_\chi(h)}\right),$$

and the claimed result follows directly from the last part of Condition H.3.

Proof of (A.8). Write

$$\mathrm{Var}(T_{1,n}) = \frac{1}{n^4\varphi_\chi^4(h)}\sum_{j=1}^{n}\left(\mathrm{E}A_j^2 - (\mathrm{E}A_j)^2\right),$$

with

$$A_j = \int K^2\left(\frac{d(\chi, \mathcal{X}_j)}{h}\right)\left(\frac{1}{g}K_0\left(\frac{y - Y_j}{g}\right) - f(y/\chi)\right)^2 w(y)dy.$$

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

Using Lemma B (noting that $K^2 = K$) we have that

$$\mathrm{E}A_j = C\frac{\varphi_\chi(h)}{g} + o\left(\frac{\varphi_\chi(h)}{g}\right). \tag{A.11}$$

Now, let us calculate the moment of order two of A_j. By conditioning on $\mathcal{X}_j$ we arrive at

$$\mathrm{E}A_j^2 = \varphi_\chi(h)\int\int\int\left(\frac{1}{g}K_0\left(\frac{y-u}{g}\right) - f(y/\chi)\right)^2 \times\left(\frac{1}{g}K_0\left(\frac{y'-u}{g}\right) - f(y'/\chi)\right)^2 f(u/\chi)w(y)w(y')dydy'du.$$

By proceeding as in the proof of Lemma B, but with a two-dimensional integration by substitution, we arrive at

$$\mathrm{E}A_j^2 = C\frac{\varphi_\chi(h)}{g^2} + o\left(\frac{\varphi_\chi(h)}{g^2}\right). \tag{A.12}$$

By using (A.11) and (A.12) we arrive at

$$\mathrm{Var}(T_{1,n}) = O\left(\frac{1}{n^3\varphi_\chi(h)^3 g^2}\right),$$

and the claimed result follows from the last part of Condition H.3.

Proof of (A.10). Let us write

$$\mathrm{E}T_{2,n} = \frac{1}{n^2\varphi_\chi^2(h)}\int\sum\sum_{i\neq j}\mathrm{E}B_i(y)B_j(y)w(y)dy,$$

with

$$B_i(y) = K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)\left(\frac{1}{g}K_0\left(\frac{y-Y_i}{g}\right) - f(y/\chi)\right), \qquad \forall i = 1,\ldots n.$$

From Lemma A, we have directly

$$\mathrm{E}\left(T_{2,n}\right) = \frac{n^2-n}{n^2\varphi_\chi^2(h)}\int(\mathrm{E}B_1(y))^2 w(y)dy = g^4 B + O(g^6),$$

and the claimed result follows from the third part of Condition H.3.

Proof of (A.9). In a first attempt, we have to compute $\mathrm{Var}(T_{2,n})$. For that, write

$$\mathrm{E}T_{2,n}^2 = \frac{1}{n^4\varphi_\chi^4(h)}\sum_{i_1\neq j_1}\sum_{i_2\neq j_2} D(i_1,j_1,i_2,j_2), \tag{A.13}$$

where

$$D(i_1,j_1,i_2,j_2) = \mathrm{E}\left(\int\int B_{i_1}(y)B_{j_1}(y)B_{i_2}(y')B_{j_2}(y')w(y)w(y')dydy'\right).$$

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

The idea is to compute the quantity $D(i_1, j_1, i_2, j_2)$ in different ways according to the number of distinct elements in the set $\{i_1, j_1, i_2, j_2\}$. Let us start with the situation when $i_1 = i_2$ and $j_1 \neq j_2$, and note that by using Lemma A and then by conditioning on $\mathcal{X}_{i_1}$ we have

$$\begin{aligned}
D(i_1, j_1, i_1, j_2) &= \int\int \left\{ \mathrm{E}\left(B_{i_1}(y) B_{i_1}(y')\right) \mathrm{E}\left(B_{j_1}(y)\right) \mathrm{E}\left(B_{j_2}(y')\right) \right\} w(y) w(y') dy dy' \\
&= g^4 V_1^2 \varphi_\chi^2(h) \int\int \left\{ \mathrm{E}\left(B_{i_1}(y) B_{i_1}(y')\right) \right\} \\
&\quad \times \left(f''(y/\chi) f''(y'/\chi) + o(1) \right) w(y) w(y') dy dy' \\
&= g^4 V_1^2 \varphi_\chi^3(h) \int\int\int \left(\frac{1}{g} K_0\left(\frac{y-u}{g} \right) - f(y/\chi) \right) \\
&\quad \times \left(\frac{1}{g} K_0\left(\frac{y'-u}{g} \right) - f(y'/\chi) \right) \\
&\quad \times f''(y/\chi) w(y) f''(y'/\chi) w(y') f(u/\chi) dy dy' du + o(\varphi_\chi^3(h) g^4) \\
&= g^4 V_1^2 \varphi_\chi^3(h) \left\{ \int f''(t/\chi)^2 w^2(t) f(t/\chi) dt \right. \\
&\quad \left. - \left(\int f''(t/\chi) w(t) f(t/\chi) dt \right)^2 \right\} + o(\varphi_\chi^3(h) g^4),
\end{aligned}$$

where

$$V_1 = \frac{1}{2} \int v^2 K_0(v) dv,$$

the last equality obtained by splitting the product involved in the integral above into four terms, and by integrating by substitution to compute each of them. In other words, we have proven that

$$D(i_1, j_1, i_1, j_2) = \frac{V}{4} \varphi_\chi^3(h) g^4 + o(\varphi_\chi^3(h) g^4), \qquad \forall j_1 \neq j_2. \tag{A.14}$$

Let us now look at the case when $i_1 = i_2$ and $j_1 = j_2$. By conditioning on $(\mathcal{X}_{i_1}, \mathcal{X}_{j_1})$, we have

$$D(i_1, j_1, i_1, j_1) = \varphi_\chi^2(h) \int\int \mathrm{E}\left(B_{i_1}(y) B_{i_1}(y')\right) \mathrm{E}\left(B_{j_1}(y) B_{j_1}(y')\right) w(y) w(y') dy dy',$$

and then by using a three-dimensional integration by substitution we obtain

$$D(i_1, j_1, i_1, j_1) = O\left(\frac{\varphi_\chi^2(h)}{g} \right), \qquad \forall i_1 \neq j_1. \tag{A.15}$$

Finally, we get from (A.13), (A.14), and (A.15),

$$\mathrm{E}T_{2,n}^2 = \frac{1}{n^4 \varphi_\chi^4(h)} \left(4n^3 \frac{V}{4} g^4 \varphi_\chi^3(h) + O\left(n^2 \frac{\varphi_\chi^2(h)}{g} \right) \right) + (\mathrm{E}T_{2,n})^2 (1 + o(1)).$$

Because of the last part of Condition H.3, this is enough to prove that

$$\mathrm{Var}(T_{2,n}) = \frac{g^4}{n \varphi_\chi(h)} V + o\left(\frac{g^4}{n \varphi_\chi(h)} \right). \tag{A.16}$$

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

Now, we will use a central limit theorem for quadratic forms. Indeed, we are in a position to apply Theorem 5.1 of De Jong (1987) with, using De Jong's notations,

$$\sigma(n)^2 = \mathrm{Var}(T_{2,n})$$

and

$$\sigma_{i,j}^2 = \frac{1}{n^4 \varphi_\chi(h)^4} \mathrm{Var}\left(\int B_i(y) B_j(y) w(y) dy \right).$$

We get directly from Lemmas A and B that

$$\sigma_{i,j}^2 = O\left(\frac{1}{n^4 \varphi_\chi(h)^2 g^2} \right). \tag{A.17}$$

Finally, because of (A.16) and (A.17), it is easy to see that the three conditions necessary to apply De Jong's theorem are fulfilled with, using again De Jong's notations,

$$K(n) = \frac{1}{g^2 \varphi_\chi(h)},$$

since we have

$$\mu_i^2 = O\left(\frac{1}{n^2 \varphi_\chi(h)^2 g^2} \right), \qquad \forall i.$$

The result of De Jong's theorem states that

$$\frac{T_{2,n} - \mathrm{E}T_{2,n}}{\sqrt{\mathrm{Var} T_{2,n}}} \to N(0,1),$$

which is enough, because of (A.16), to get the claimed result (A.9).

Finally, (A.7), (A.8), (A.9), and (A.10) have been checked, and this proof is finished. ■

Proof of Lemma 2. Note that, under H_0, and by using Condition H.4, we have uniformly over y,

$$|f(y/\chi) - f_{\hat{\theta}_0}(y/\chi)| = O(U_n) = o\left(\frac{1}{\sqrt{n \varphi_\chi(h)}} \right), \quad \text{a.s.} \tag{A.18}$$

We will prove the three assertions of Lemma 2, by using several times the result (A.18).

Term $T_{3,n}$. By using (A.18) and Condition H.3, we get

$$\begin{aligned} \mathrm{E}(T_{3,n}) &= O\left(n^{-2} \varphi_\chi^{-2}(h) U_n^2 \mathrm{E}\left[\sum_{i=1}^n K\left(\frac{d(\chi, \mathcal{X}_i)}{h} \right) \right]^2 \right) \\ &= O\left(\frac{1}{n^2 \varphi_\chi^2(h)} U_n^2 \left(n \varphi_\chi(h) + n(n-1) \varphi_\chi^2(h) \right) \right) \\ &= o\left(\sqrt{\frac{g^4}{n \varphi_\chi(h)}} \right). \end{aligned} \tag{A.19}$$

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

We compute the variance of this term using the same arguments as before to treat the term $T_{2,n}$, and by (A.18) and Condition H.3, we get

$$\mathrm{Var}\left(T_{3,n}\right) = O\left(\frac{U_n^4}{n\varphi_\chi(h)}\right) = o\left(\frac{g^4}{n\varphi_\chi(h)}\right). \tag{A.20}$$

Finally, (A.19) and (A.20) are enough to prove Lemma 2 for this term.

Term $T_{4,n}$. Because $T_{4,n}$ has the same structure as $T_{1,n}$, it can be treated following the same steps as those for Lemma 1. The additional factor will be easily controlled by means of (A.18). For the bias part, we get from (A.18) and Condition H.3,

$$\mathrm{E}\left(T_{4,n}\right) = O\left(\frac{1}{n^4\varphi_\chi^4(h)} U_n n\varphi_\chi(h)\right) = o\left(\sqrt{\frac{g^4}{n\varphi_\chi(h)}}\right).$$

For the variance term, using (A.18) and Condition H.3, we arrive at

$$\mathrm{Var}\left(T_{4,n}\right) = O\left(\frac{U_n^2}{n^3\varphi_\chi^3(h)g}\right) = o\left(\frac{g^4}{n\varphi_\chi(h)}\right).$$

We now conclude this proof, as before for the term $T_{3,n}$.

Term $T_{5,n}$. Note that the term $T_{5,n}$ has the same structure as $T_{2,n}$, with an additional component that can be controlled by means of (A.18). For the expectation, by following the same steps as those for Lemma 1 when treating $T_{2,n}$, we obtain that

$$\mathrm{E}\left(T_{5,n}\right) \le C\left(\frac{1}{n^2\varphi_\chi^2(h)} U_n n^2\varphi_\chi^2(h)g^2\right), \tag{A.21}$$

which is enough by using (A.18) and Condition H.3 to show that

$$\mathrm{E}\left(T_{5,n}\right) = o\left(\sqrt{\frac{g^4}{n\varphi_\chi(h)}}\right).$$

Similarly, the variance term can be calculated by following the same steps as for $T_{2,n}$, but using additionally the bound (A.18), and we obtain

$$\mathrm{Var}\left(T_{5,n}\right) = O\left(\frac{U_n^2 g^2}{n\varphi_\chi(h)}\right) = o\left(\frac{g^4}{n\varphi_\chi(h)}\right).$$

We now conclude this proof, as before for the term $T_{3,n}$.

The proof of Lemma 2 is now complete. ■

Proof of Lemma 3.

Proof of (A.4). This term needs special attention, since it does not have the same behavior under the null as under the alternative hypothesis. The difference in behavior can be stated by using the following result, which can be derived directly from the definition of H_1 and by using Condition H.5:

$$\left|f(y/\chi) - f_{\hat{\theta}_0}(y/\chi)\right| = \varepsilon_n s(y) + o(\varepsilon_n), \quad \text{a.s.}, \tag{A.22}$$

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press

where the o is uniform over $y \in S_w$. If we compute the bias of this term as we did in Lemma 1 for the bias of $T_{2,n}$, and by using the result (A.22), we get

$$\mathrm{E}\left(T_{5,n}\right) = C\varepsilon_n g^2 \int f''(y/\chi)s(y)w(y)dy + o(\varepsilon_n g^2).$$

We directly conclude the result by using Condition H.5.

Proof of (A.5). Following the same lines as with Lemma 1 for treating the term $T_{2,n}$, but using (A.22), we arrive directly at

$$\mathrm{Var}\left(T_{5,n}\right) = O\left(\frac{g^2\varepsilon_n^2}{n\varphi_\chi(h)}\right).$$

Combined with Condition H.5 and (A.4), this last result is enough to get (A.5).

Proof of (A.6) for $T_{3,n}$. Following the same lines as with Lemma 2, but using (A.22) instead of (A.18), the bias term can be computed as

$$\begin{aligned}\mathrm{E}\left(T_{3,n}\right) &= O\left(n^{-2}\varphi_\chi^{-2}(h)\varepsilon_n^2\mathrm{E}\left[\sum_{i=1}^{n} K\left(\frac{d(\chi,\mathcal{X}_i)}{h}\right)\right]^2\right)\\ &= O\left(\frac{1}{n^2\varphi_\chi^2(h)}\varepsilon_n^2\left(n\varphi_\chi(h)+n(n-1)\varphi_\chi^2(h)\right)\right),\end{aligned}$$

which, because of (A.22), leads to

$$\mathrm{E}\left(T_{3,n}\right) = o\left(\mathrm{E}\left(T_{5,n}\right)\right). \tag{A.23}$$

Similarly, we have for the variance term that

$$\mathrm{Var}\left(T_{3,n}\right) = O\left(\frac{\varepsilon_n^4}{n\varphi_\chi(h)}\right) = o\left(\left(\mathrm{E}\left(T_{5,n}\right)\right)^2\right). \tag{A.24}$$

The results (A.23) and (A.24) complete the proof of (A.6) for this term.

Proof of (A.6) for $T_{4,n}$. The proof is the same as that for Lemma 2 but using (A.22) instead of (A.18), and we have the results

$$\mathrm{E}\left(T_{4,n}\right) = o\left(\mathrm{E}\left(T_{5,n}\right)\right),$$

and

$$\mathrm{Var}\left(T_{4,n}\right) = o\left(\left(\mathrm{E}\left(T_{5,n}\right)\right)^2\right),$$

which are enough to get the proof of (A.6) for this term.

The proof of Lemma 3 is now complete. ■

https://doi.org/10.1017/S0266466611000351 Published online by Cambridge University Press