View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Columbia University Academic Commons

Columbia University

Department of Economics
Discussion Paper Series

Nonparametric Specification Analysis of Dynamic Parametric Models

Marc Henry
Olivier Scaillet

Discussion Paper #:0102-20

Department of Economics
Columbia University
New York, NY 10027

March 2002

Nonparametric specification analysis of dynamic parametric models[1]

Marc Henry[†*] and Olivier Scaillet[*]

[†]Columbia University in the City of New York
[*]Université Catholique de Louvain

First version: 28 August 2000

Abstract

Time series parametric models generally cater to a particular objective, such as forecasting, and it is therefore desirable to judge such models solely on the basis of their performance in the fullfillment of that objective. We propose a specification testing procedure which concentrates power on the parametric model's ability to estimate a set of characteristics of the finite dimensional distributions of the process. It is based on the comparison between a nonparametric estimate of the said characteristic and its parametric bootstrap analogue. Applications of this principle are proposed for the assessment of recursive dynamic models in the estimation of conditional means and conditional quantiles for mixing processes and for the estimation of dependence in long memory processes.

1 Introduction

The various specification testing procedures for likelihood models unified under the m-test framework of Newey (1985) and Tauchen (1985), such as the Lagrange multiplier specification test, the Hausman test, Cox's test of non-nested hypotheses, Newey's conditional moments test and White's information matrix test, are parametric in nature, and can fail to have power against certain departures from the null hypothesis of correct specification.

[1]The authors would like to thank Xiaohong Chen, Frank Diebold, Christian Gouriéroux, Peter M. Robinson, Frank Schorfheide, Jean-Michel Zakoïan and seminar participants at CORE, LSE and UPenn for helpful discussions. This research was carried out while the first author was visiting UCL, and financial support from the latter institution is gratefully acknowledged. The second author also acknowledges financial support from SSTC grant PAI number P4/01. Part of this research was carried out while he was visiting THEMA. The usual disclaimer applies. Correspondance address: marc.henry@columbia.edu

This is remedied by the use of testing principles based on distances between non-parametric and parametric counterparts and inspired by the Kolmogorov-Smirnov and Cramér-von Mises tests. Some relevant cases are Eubank and Spiegelman (1990), Wooldridge (1992), Haerdle and Mammen (1993), Gozalo (1993), and Zheng (1996). This "nonparametric" approach was mainly developped within the framework of independently and identically distributed random variables to assess the choice of functional form for regressions or the choice of parametric conditional density models. In the former case, parametric and nonparametric estimates of the conditional mean are compared; in the latter, parametric density estimates are compared to nonparametric ones. Stintchcombe and White (1998) show that these tests and other consistent tests for arbitrary misspecification, such as Bierens's (Bierens (1990)), based on the nuisance parameter approach, and Robinson's entroby-based testing procedure (Robinson (1991)), are all derived from estimates of distances with the relevant choice of topology.

Here we consider a related procedure designed to test the specification of time series models designed to achieve a particular objective (such as forecasting through conditional means or conditional medians) for which consistent estimation of only a finite number of characteristics of the probability model (such as conditional moments, conditional distribution quantiles, local properties of the spectral density, etc...) is relevant. We therefore propose to concentrate the power of the test on the estimation of such characteristics by evaluating a distance between some non-parametric estimate of the said characteristic and a commensurate estimate of its parametric bootstrap analogue. This can be easily achieved (however complicated the model, but particularly in case the specified model is Markovian) by a comparison between the nonparametric estimate based on the originial data, and the same nonparametric estimate based on a parametric bootstrap of the original data.

The procedure is inspired by the encompassing principle applied to non nested hypotheses testing (see Mizon and Richard (1986), Gouriéroux and Monfort (1995) and Dhaene, Gouriéroux, and Scaillet (1998)) and the simulation based indirect inference method developed in Gouriéroux, Monfort, and Renault (1993). This procedure departs from the above in that it does not require the specification of a rival parametric model, but only requires a consistent parametric estimate of the true or some "pseudo-true" parameter value under the null (in the terminology of White (1982) and Gouriéroux, Monfort, and Trognon (1984)), and the availability of a suitable nonparametric estimate of the relevant feature, from which the power properties of the test are derived.

The paper is organised as follows. In Section 2 we decribe the specification testing principle in general terms, whereas section 3 considers a specialization of the principle to the assessment of the adequacy of a recursive dynamic model to estimating characteristics of the conditional distribution, such as conditional mean and quantiles. In particular we examine joint tests of conditional moments and conditional quantiles. This is for example relevant for testing the restriction that conditional mean equal conditional median induced by symmetric innovations in nonlinear parametric regression models. Section 4 considers a specialization of the principle to the use of parametric models of the spectral density function of long range dependent processes. This is particularly relevant for the estimation of dependence in the form of the long memory parameter. Proofs and mathematical developments are gathered in an appendix.

2 Nonparametric specification testing

Consider a stochastic process $\{Y_t, t \in \mathbb{Z}\}$ on a probability space $(\Omega, \mathcal{F}, P_0)$ and assume our data set consists in a realization $Y = (Y_1, \ldots, Y_T)'$ from that process. Let $\mathcal{G}(.)$ be an $\mathbb{R}^d$-valued characteristic of the finite-dimensional distributions of the process, with unknown true value $\mathcal{G}_0 = \mathcal{G}(P_0)$. We wish to assess the suitability of a particular class of models $M = \{\mu_\theta, \theta \in \Theta \subset \mathbb{R}^q\}$ for the estimation of $\mathcal{G}_0$, and we define the parameters of interest by a mapping $\mathcal{V} : M \to \Theta$. Let θ_0 be the true value of the parameter vector under correct specification of the model class, i.e. $P_0 \in M$, in which case $\theta_0 = \mathcal{V}(P_0)$.

Consider a sequence $\hat{\theta}_T$, which may be a sequence of extremum estimates[2], and converges in probability to a value $\theta_0 \in \mathrm{int}(\Theta)$ which satisfies $\theta_0 = \mathcal{V}(P_0)$ if $P_0 \in M$.

The null hypothesis to be tested is therefore

$$\mathrm{H}_0 : \quad \mathcal{G}_0 = \mathcal{G}(\mu_{\theta_0}) \tag{1}$$

against the alternative

$$\mathrm{H}_a : \quad \mathcal{G}_0 \neq \mathcal{G}(\mu_{\theta_0}). \tag{2}$$

We are therefore not concerned with the "true" specification, and for our purposes, the true probability measure P_0 for the stochastic process may be outside the class M of models for its set of finite dimensional distributions, as long as the value for $\mathcal{G}(\mu)$ implied by the use of M is correct.

Suppose in addition that under H_0,

$$\sqrt{T}(\hat{\theta}_T - \theta_0) = O_p(1), \tag{3}$$

[2] If model class M is a family generating likelihoods and $\hat{\theta}$ is a direct or indirect pseudo maximum likelihood estimator sequence, or if model class M is a family generated by moment functions, and $\hat{\theta}$ is a GMM estimator (see White (1987) for definitions), then Domowitz and White (1982) and Gouriéroux, Monfort, and Trognon (1984) give very general conditions under which $\hat{\theta}_T$ converges a.s.-P_0 to a sequence θ_T^*, but the latter may not converge. In the iid case, Huber (1967) and White (1982) give conditions under which a pseudo maximum likelihood estimator sequence converges at rate $\sqrt{T}$ to a limit θ_0, which is the set of parameters that maximize the Kullback-Leibler Information Criterion.

and consider a $\mathcal{C}^1$ nonparametric estimator for $\mathcal{G}$, denoted $\tilde{\mathcal{G}}(Y)$ such that

$$\sqrt{Th}(\tilde{\mathcal{G}}(Y) - \mathcal{G}_0) \rightarrow_d N(0, V_0) \tag{4}$$

where h is generically called the bandwidth and satisfies

$$h + (Th)^{-1} \rightarrow 0 \quad \text{as} \quad T \rightarrow \infty. \tag{5}$$

Our specification testing procedure is based on the following principle: under H_0, the nonparametric estimator will have the same asymptotic properties whether it is based on the sample of real data or on a sample simulated from the estimated model $\mu_{\hat{\theta}}$ for the finite dimensional distributions of the process. Call $Y(\theta)$ a sample simulated from μ_θ, we have, under the same regularity conditions as for (4):

$$\sqrt{Th}(\tilde{\mathcal{G}}(Y(\theta)) - \mathcal{G}(\mu_\theta)) \rightarrow_d N(0, V_\theta) \tag{6}$$

With N conditionnally independent samples $Y^s(\hat{\theta})$, $s = 1$ to N, simulated from $\mu_{\hat{\theta}}$, we can form the Wald vector $\hat{\omega} = \tilde{\mathcal{G}}(Y) - \frac{1}{N}\sum_{s=1}^{N} \tilde{\mathcal{G}}(Y^s(\hat{\theta}_T))$ and the test statistic $\xi = Th\hat{\omega}'\hat{\Sigma}^\dagger\hat{\omega}$ is asymptotically $\mathcal{X}^2$ distributed under H_0, if $\hat{\Sigma}^\dagger$ is a consistent estimator of the inverse of the asymptotic variance of $\hat{\omega}$.

In the general case, one may forego the estimation of $\Sigma^\dagger$ whose principle may differ for different choices of criterion $\mathcal{G}$, through the use of the asymptotic in N using the following result:

<u>Proposition 1:</u>

Putting $N = N(T)$ and supposing that $N(T) \rightarrow \infty$ when $T \rightarrow \infty$, we have, under (3) and (4): $Th\hat{\omega}'\tilde{V}^\dagger\hat{\omega} \rightarrow_a \mathcal{X}_d^2$, where $\tilde{V}^\dagger$ is a consistent estimate of V_0^{-1}.

Consider local alternatives of the form $\mathcal{G}_T = \mathcal{G}(\mu_{\theta_0}) + \Delta_T$. Under such local alternatives,

$$\sqrt{Th}\hat{\omega} = \sqrt{Th}(\tilde{\mathcal{G}}(Y) - \mathcal{G}_T) + \frac{\sqrt{Th}}{N}\sum_{s=1}^{N}(\tilde{\mathcal{G}}(Y^s(\hat{\theta})) - \mathcal{G}(\mu_{\theta_0})) + \sqrt{Th}\Delta_T,$$

so that the test is sensitive to all deviations from the hypothesis of order $(Th)^{-\frac{1}{2}}$. Therefore, under additional regularity conditions, the procedure can have power against alternatives approaching the null at a rate arbitrarily close to $T^{-\frac{1}{2}}$ with the use of a kernel of suitably high order.

Of course the Monte Carlo approximation of the parametric bootstrap finite dimensional distributions is particularly suitable in the assessment of recursive models, from which it is easy to draw simulated data samples. This particular case will be considered in the next section. However, simulating samples from non recursive models such as models for long range dependent time series may be quite arduous and introduce additional errors. To circumvent this, a frequency domain strategy is proposed for estimates based on functions of periodogram ordinates: the test is the based on the comparison between $\tilde{\mathcal{G}}(Y)$ and its "parametric analogue" where periodogram ordinates are replaced by ordinates of the parametrically estimated spectral density.

Note that in case of misspecification of model class M, θ_0 may still be defined, as the probability limit of the sequence $\hat{\theta}_T$, supposed convergent, so that H_0 may still hold, and the testing procedure we consider is a joint test of the suitability of both model class M and the parametric method employed to compute $\hat{\theta}_T$ for the estimation of $\mathcal{G}_0$.

3 Conditional expectations and quantiles

In this first instance, we wish to analyse the following recursive model for a strictly stationary time series:

$$Y_t = r(X_t, \epsilon_t; \theta), \qquad \theta \in \Theta \subset \mathbb{R}^q, \tag{7}$$

where r is a known function, $X_t = (Y_{t-1}, ..., Y_{t-p})'$, and $\{\epsilon_t\}$ is a sequence of i.i.d. innovations with known distribution. Since ϵ_t has a known distribution it is in principle possible to derive the conditional p.d.f. of Y_t given X_t denoted by $f(y|x;\theta)$, as the p.d.f. of the image of the distribution of ϵ_t by $r(X_t, \cdot; \theta)$. Hence Equation (7) implicitly defined a parametrized family of density functions $f(y|x;\theta)$. No matter how complicated $f(y|x;\theta)$, it is easy to draw from it, because of the recursive form of the model.

In the following we assume that we are provided an estimator $\hat{\theta}$ of the parameter θ, satisfying condition (3),[3] and we may simulate N paths of length T : $\{Y_t^s(\hat{\theta}); t = 1, ..., T; s = 1, ..., N\}$ using Equation (7).

As explained in the previous section, we do not intend to verify whether the true conditional distribution belongs to the model implied from (7). We are only interested in weaker constraints induced by Equation (7) on the data generating process. These constraints concern conditional characteristics of the data.

Let us take a positive integer n, and let $0 < \tau_1 < ... < \tau_n$ be integers, so that we define $Z_t = (Y_{t-\tau_1}, ..., Y_{t-\tau_n})'$. In particular we may take $n = p$, $\tau_1 = 1$, $\tau_n = p$, which gives : $Z_t = X_t$. We denote by $f_0(y, z)$, $F_0(y, z)$, the marginal p.d.f. and c.d.f. of $(Y_t, Z_t')'$, while the conditional p.d.f. and c.d.f. are written $f_0(y|z)$, and $F_0(y|z)$,

[3]This estimator may correspond either to a direct estimator obtained by a pseudo maximum likelihood method or an indirect estimator obtained by a simulation based method. In the former θ_0 is called a pseudo-true value while in the latter it is called an indirect pseudo-true value (see Gouriéroux and Monfort (1997)).

respectively. The corresponding p.d.f. and c.d.f. implied from (7) will have θ as argument instead of being subscripted by 0.

Choosing z as a conditioning point of interest, the testing hypothesis (1) can be refined to testing for equality between the conditional expectation of Y under $f_0(y|z)$ and $f(y|z;\theta_0)$ in the form:

$$\mathrm{H}_0: \quad E_0\left[g(Y)|z\right] = E_{\theta_0}\left[g(Y)|z\right],$$

or similarly for conditional quantiles, in the form:

$$\mathrm{H}_0: \quad \inf_{y\in\mathbb{R}} \{y : F_0(y|z) \geq p\} = \inf_{y\in\mathbb{R}} \{y : F(y|z;\theta_0) \geq p\}$$

We now need to estimate conditional moments and conditional quantiles of the distribution of Y_t nonparametrically.

For conditional moments, we look at the quantities

$$G(\zeta_i)f(\zeta_i) \equiv E(g(Y_t)|Z_t = \zeta_i)f(\zeta_i) \tag{8}$$

at distinct points $\zeta_i \in \mathbb{R}^n$, $i = 1,\ldots,d$, the pdf of Z_t is supposed to exist and is denoted f, and g is a Borel function on $\mathbb{R}$ such that $E|g(Y_t)| < \infty$.

For a quantile of order $p \in (0,1)$, we assume that the cumulative distribution function $F(.|\zeta_i)$ of Y_t given Z_t at distinct points ζ_i is such that the equation $F(y|\zeta_i) = p$ admits a unique solution for each of the ζ_i denoted $Q(\zeta_i,p)$.

Let $k_{ij}(u)$ be a real bounded and symmetric function on $\mathbb{R}$ such that

$$\int k_{ij}(u)du = 1, \quad i = 1,\ldots,d, \quad j = 1,\ldots,n,$$

and

$$K_i(u;h^{(i)}) = \Pi_{j=1}^n k_{ij}(u_j/h_{ij}), \quad i = 1,\ldots,d,$$

where $h^{(i)}$ is the diagonal matrix with diagonal elements $(h_{ij})_{j=1}^n$ and the bandwidths h_{ij} are positive functions of T such that

$$|h^{(i)}| + (T|h^{(i)}|)^{-1} \to 0 \ \text{ when } \ T \to \infty.$$

In addition, let l and h or l_i and h_i, $i = 1,\ldots,d$, satisfy the same conditions as any of the k_{ij} and h_{ij}.

The conditional expectation estimator will be based on the following estimate of (8):

$$[g;\zeta_i] = (T|h^{(i)}|)^{-1} \sum_{t=1+\tau_n}^{T} g(Y_t)K_i(\zeta_i - Z_t;h^{(i)}).$$

The pdf of (Z_t) at (ζ_i), denoted $f(\zeta_i)$, will be estimated by

$$[1;\zeta_i] = (T|h^{(i)}|)^{-1} \sum_{t=1+\tau_n}^{T} K_i(\zeta_i - Z_t; h^{(i)}),$$

so that the conditional expectation of $g(Y_t)$ given $Z_t = \zeta_i$ is estimated by

$$\hat{G}(\zeta_i) = [g;\zeta_i]/[1;\zeta_i].$$

The pdf of (Y_t, Z_t) at (ξ_i, ζ_j), denoted $f(\xi_i, \zeta_j)$, will be estimated by

$$[1;\xi_i,\zeta_j] = (Th_i|h^{(j)}|)^{-1} \sum_{t=1+\tau_n}^{T} l_i(h_i^{-1}(\xi_i - Y_t))K_j(\zeta_j - Z_t; h^{(j)}).$$

Finally, the conditional cumulative distribution of Y_t given $Z_t = \zeta_j$ will be estimated at distinct points ξ_i, $i = 1, \ldots, d$, by

$$\hat{F}(\xi_i|\zeta_j) = \int_{-\infty}^{\xi_i} [1;u,\zeta_j]du/[1;\zeta_j] \equiv \hat{\phi}(\xi_i,\zeta_j)/[1;\zeta_j],$$

(calling $\phi(\xi_i, \zeta_j) = \int_{-\infty}^{\xi_i} f(u, \zeta_j)du$) and the conditional quantile $Q(\zeta_j, p_i)$ will be estimated by

$$\hat{Q}(\zeta_j, p_i) = \inf_{y \in \mathbb{R}} \left\{ y : \hat{F}(y|\zeta_j) \geq p_i \right\}.$$

In addition, call $f(\xi|\zeta_i)$ and $\hat{f}(\xi|\zeta_i)$ the first derivatives with respect to ξ of $F(\xi|\zeta_i)$ and $\hat{F}(\xi|\zeta_i)$.

3.1 Strong mixing conditions

Let $\mathcal{M}_a^b$ be the σ-field of events generated by Y_t, $a \leq t \leq b$, and introduce the Rosenblatt strong-mixing coefficients

$$\alpha_j = \sup_{A \in \mathcal{M}_{-\infty}^t, B \in \mathcal{M}_{t+j}^\infty} |P(A \cap B) - P(A)P(B)|, \quad j > 0.$$

We assume that there exists some $\theta > 2$ such that

$$E|g(Y_t)|^\theta < \infty \text{ and } \sum_{j=N}^{\infty} \alpha_j^{1-2/\theta} = O(N^{-1}),\ N \to \infty$$

3.2 Conditions on densities and kernels

(i) For all $i = 1, \ldots, d$ we have $f(\zeta_i) > 0$.

(ii) Continuous second order partial derivatives for the pdf of (Y_t, Z_t) in neighbourhoods of all the pairs (ξ_i, ζ_j) where estimation is performed.

(iii) Bandwidths satisfying $|h^{(i)}|||h^{(i)}||^4 T \to 0$.

(iv) $|k_{ij}(u)| \leq C(1 + |u|)^{-(1+\omega_i/n)}$, and $||h^{(i)}||^{n+\omega_i-2} \leq C|h^{(i)}|$, $\omega_i > 2$.

(v) The pdf of (Z_t, Z_{t+s}) exists and is bounded in a neighbourhood of all pairs (ζ_i, ζ_j), $i, j = 1, \ldots, d$, uniformely in $s > 1$.

3.3 Theorem 1: Asymptotic normality of the conditional expectation estimator

Let S be the d dimensional vector with components $S_i/\hat{V}_i$ with:

$$S_i = (T|h^{(i)}|)^{1/2}\left\{\hat{G}(\zeta_i) - G(\zeta_i)\right\}$$

$$\hat{V}_i^2 = \int \frac{\Pi_{j=1}^n k_{ij}^2(u)du_j}{[1;\zeta_i]^2}\left\{[g^2;\zeta_i] - \frac{[g;\zeta_i]^2}{[1;\zeta_i]}\right\}$$

Under the strong mixing conditions, the conditions on densities and kernels above and the additional conditions that G is twice continuously differentiable at all ζ_i, that $E(g^2(Y_t)|Z_t = z)$ is continuous at all ζ_i and that for some $\gamma > \theta$, $E(|g(Y_t)|^\gamma|Z_t = z)$ is bounded in the neighbourhood of each of the ζ_i, $i = 1, \ldots, d$, S converges in distribution to a vector of independent standard normal random variables.

3.4 Theorem 2: Asymptotic normality of the conditional quantile estimator

Consider a single level p for the quantile. Let S be the d dimensional vector with components $S_i/\hat{V}_i$ with:

$$S_i = (T|h^{(i)}|)^{1/2}\left\{\hat{Q}(\zeta_i, p) - Q(\zeta_i, p)\right\}$$

$$\hat{V}_i^2 = \frac{p(1-p)}{[1;\zeta_i]}\left\{\hat{f}(\hat{Q}(\zeta_i, p)|\zeta_i)\right\}^{-2}\int \Pi_{j=1}^n k_{ij}^2(u)du_j$$

Theorem

Under the additional condition that for each ζ_i, $f(Q(\zeta_i, p)|\zeta_i) > 0$, S converges in distribution to a vector of independent standard normal random variables.

3.5 Theorem 3: Specification test statistic

Consider N independent simulated samples from the estimated distribution $f(y|z,\hat{\theta})$ and denote by $\hat{Q}^s(\zeta_i,p)$ the analogue of the nonparametric estimator of, say, the conditional quantile of level p at point ζ_i based on sample $Y^s(\hat{\theta})$. Under the conditions of Theorem 2, calling ω_0 is the vector with components

$$\omega_i = \inf_{y\in\mathbb{R}} \{y : F_0(y|\zeta_i) \geq p\} - \inf_{y\in\mathbb{R}} \{y : F(y|\zeta_i;\theta_0) \geq p\},$$

and V_0 and $V(\theta)$ the diagonal matrices with diagonal elements, respectively

$$V_{i0}^2 = \frac{p(1-p)}{f_0(\zeta_i)} \{f_0(Q(\zeta_i,p)|\zeta_i)\}^{-2} \int \Pi_{j=1}^n k_{ij}^2(u)du_j$$

and

$$V_i^2(\theta) = \frac{p(1-p)}{f(\zeta_i;\theta)} \{f(Q(\zeta_i,p;\theta)|\zeta_i;\theta)\}^{-2} \int \Pi_{j=1}^n k_{ij}^2(u)du_j$$

the Wald vector ω with components

$$\hat{\omega}_i = \hat{Q}(\zeta_i,p) - \frac{1}{N}\sum_{s=1}^{N} \hat{Q}^s(\zeta_i,p)$$

has the following asymptotic distributions:

- Under H_0: $(TH)^{1/2}\hat{\omega} \rightarrow_d N(0,\Sigma)$
- Under H_a: $(TH)^{1/2}(\hat{\omega}-\omega_0) \rightarrow_d N(0,\Sigma)$,

with $\Sigma = \left(V_0 + \frac{1}{N}V(\theta_0)\right)^{-1}$, H is the diagonal matrix with diagonal elements $|h^{(i)}|$, from which we can derive the properties of the test statistic $\hat{\xi} = TH\hat{\omega}'\hat{\Sigma}\hat{\omega}$, with $\hat{\Sigma}$ a consistent estimate of Σ, and establish consistency and local power of the testing procedure[4].

4 Valid local spectral density estimation with misspecified models

In this section, we consider the estimation of the spectral density matrix of a bivariate process in a preselected band of frequencies of interest. The focus on the

[4]The diagonal matrix with diagonal elements $\left(\hat{V}_i^2 + \frac{1}{N}V^2(\hat{\theta})\right)^{-1}$ is a candidate.

bivariate does not entail any real loss of generality but greatly simplifies notation. Consider Y_t to be a bivariate generalized linear process

$$Y_t - \mu_Y = \sum_{j=0}^{\infty} \pi_j \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} ||\pi_j||^2 < \infty, \tag{9}$$

where μ_Y is the mean of the process, $||.||$ is a matricial norm and ε_t is a bivariate matringale difference process.

let $f_Y(\lambda)$, $\lambda \in [0, \pi)$, be the spectral density matrix of the process Y_t. The process is supposed to be covariance stationary, but unlike in section 3, we do not assume weak dependence via mixing conditions. Instead, we allow the process to have long memory (resp. seasonal/cyclical long memory) characterized by the existence of hyperbolic singularities in the spectral density at zero frequency (resp. some nonzero frequency).

In certain contexts, one may choose to investigate the dynamics of such series in a band of frequencies bounded away from potential singularities, using misspecified Markov models either because they are structural, or because they are simpler to implement.

In such a setting, mild local regularity conditions on the spectral density matrix in addition to integrability (imposed by covariance stationarity) would consitute the nonparametric framework to serve as a benchmark for the local specification analysis.

We therefore consider two nested specifications, one of which we call "nonparametric specification," for which we have an asymptotically normal estimator of the spectral density matrix at all regularity points of the spectrum, and one which we call "parametric specification" which could typically be a stationary vector autoregression with normally and identically distributed innovations.

We define the "nonparametric specification" with the following set of assumptions:

- **N1:** $f_Y(\lambda)$ is twice continuously differentiable on $[\lambda_1, \lambda_2]$, for some $0 < \lambda_1 < \lambda_2 < \pi$.

- **N2:** (9) holds such that the innovations have finite fourth moments, fixed conditional second moments and fourth cummulants satisfying

$$\max_u |\kappa_u(\alpha, \beta, \gamma, \delta)| < \infty \quad \text{for all} \quad \alpha, \beta, \gamma, \delta = 1, 2.$$

- **N3:** The diagonal elements of the autocovariance function γ_j^a, $a = 1, 2$ satisfy

$$\gamma_j^a = O(j^{2\zeta_a - 1}), \quad \zeta_a < \frac{1}{2}, \quad a = 1, 2.$$

The nonparametric specification covers specific models such as fractional Autoregressive moving average (ARFIMA), cyclical and seasonal models such as cyclical ARFIMA and Gegenbauer ARMA (as in Gray, Zhang, and Woodward (1989)).

Letting

$$d_Y(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} (Y_t - \bar{Y}_t) e^{-it\lambda}$$

be the discrete Fourier transform, and $I_Y(\lambda) = d_Y(\lambda) d_Y^*(\lambda)$ be the periodogram of the process, the smoothed periodogram estimator

$$\hat{f}_Y(\lambda) = \int_{-\pi}^{\pi} mW(m(\lambda - \omega)) I_Y(\omega) d\omega,$$

is consistent and asymptotically normal for $f_Y(\lambda)$ for any $\lambda \in [\lambda_1, \lambda_2]$ under the nonparametric specification and the following assumptions on the spectral window W and the bandwidth m (see for instance Hidalgo (1996)):

- **B1:** $Tm^{-5} + m^3 T^{-1} \log^2 T \to 0, \quad \text{as} \quad T \to \infty.$
- **W1:** W is even, positive and twice continuously differentiable and satisfies

$$\int W(x) dx = 1 \quad \text{and} \quad \int x^2 W(x) dx < \infty.$$

The validation procedure presented in this paper is particularly suited to this context as the use of a misspecified structural stationary vector autoregressive representation to investigate the dynamics of the process on the $[\lambda_1, \lambda_2]$ frequency range may produce inconsistent estimates of the spectral density matrix, but such inconsistency would be picked up by the validation test. In the same way as the test in section 3 was adapted to a collection of conditional mean and quantiles, we can now check whether the chosen markovian misspecified model produces consistent estimates of coherence, phase, frequency response and other functionals of the spectral density matrix of the bivariate process[5].

If one is interested, for instance, in phase and coherence estimation at a collection of frequencies $\omega_1, \ldots, \omega_d$ such that $\lambda_1 < \omega_i < \lambda_2$, $i = 1, \ldots, d$, one may naturally

[5]It should be noted that Diebold, Onahan, and Berkowitz (1998) propose an empirical test of second order adequacy between model and data which is in some way the dual of our procedure as they compare the model spectral density with Bonferroni confidence tunnels based on bootstrap replications of the data spectral density.

base the validation Wald test statistic on the following corrolary to Theorem 2.1 in Hidalgo (1996):

Theorem 4:

Under the "nonparametric specification," the estimators for coherence $C(\lambda)$ and phase $\phi(\lambda)$ defined by

$$\hat{C}(\lambda) = \frac{|\hat{f}_Y^{12}|}{\sqrt{\hat{f}_Y^{11}\hat{f}_Y^{22}}} \quad \text{and} \quad \hat{\phi}(\lambda) = \tan^{-1}\frac{\mathcal{I}m(\hat{f}_Y^{12})}{\mathcal{R}e(\hat{f}_Y^{12})}, \quad |C(\lambda)| \neq 0$$

and $\hat{C} = (\hat{C}(\omega_1), \ldots, \hat{C}(\omega_d))$, $\hat{\phi} = (\hat{\phi}(\omega_1), \ldots, \hat{\phi}(\omega_d))$, satisfy

$$m^{-1/2}T^{1/2}(\hat{C} - C) \rightarrow N(0, \text{diag}(1 - |C(\omega_i)|^2)_{i=1}^d)$$

$$m^{-1/2}T^{1/2}(\hat{\phi} - \phi) \rightarrow N(0, \text{diag}(|C(\omega_i)|^{-2} - 1)_{i=1}^d),$$

and the asymptotic variances can be replaced by sample analogues in the usual way.

References

Berlinet, A., A. Gannoun, and E. Matzner-Lober (1998): "Propriétés asymptotiques d'estimateurs convergents de quantiles conditionnels," *Comptes Rendus de l'Académie des Sciences*, 326, 611–614.

Bierens, H. (1990): "A consistent conditional moment test for functional form," *Econometrica*, 58, 1443–1458.

Dhaene, G., C. Gouriéroux, and O. Scaillet (1998): "Indirect encompassing and instrumental models," *Econometrica*, 66, 673–688.

Diebold, F., Onahan, and Berkowitz (1998): "Dynamic equilibrium economies: a framework for comparing models and data," *Review of Economic Studies*, 65, 431–451.

Domowitz, I., and H. White (1982): "Misspecified models with dependent observations," *Journal of Econometrics*, 20, 35–58.

Eubank, R., and P. Spiegelman (1990): "Testing the goodness of fit of a linear model via nonparametric regression techniques," *Journal of the American Statistical Association*, 85, 387–392.

Gouriéroux, C., and A. Monfort (1995): "Testing, encompassing and simulating dynamic econometric models," *Econometric Theory*, 11, 195–228.

GOURIÉROUX, C., and A. MONFORT (1997): *Simulation based econometric methods*, CORE Lecture Series. Oxford University Press: Oxford.

GOURIÉROUX, C., A. MONFORT, and E. RENAULT (1993): "Indirect inference," *Journal of Applied Econometrics*, 8, S85–S118.

GOURIÉROUX, C., A. MONFORT, and A. TROGNON (1984): "Pseudo maximum likelihood methods: Theory," *Econometrica*, 52, 681–700.

GOZALO, P. (1993): "A consistent model specification test for nonparametric estimation of regression function models," *Econometric Theory*, 9, 451–477.

GRAY, H. L., N. F. ZHANG, and W. A. WOODWARD (1989): "On generalized fractional processes," *Journal of Time Series Analysis*, 10, 233–257.

HAERDLE, W., and E. MAMMEN (1993): "Comparing nonparametric versus parametric regression fits," *Annals of Statistics*, 21, 1926–1947.

HIDALGO, J. (1996): "Spectral analysis for bivariate time series with long memory," *Econometric Theory*, 12, 773–792.

HUBER, P. (1967): "The behaviour of maximum likelihood estimators under nonstandard conditions," in *Fifth Berkeley Symposium in Mathematical Statistics and Probability*, ed. by J. Neyman, and J.-M. LeCam, vol. 1, pp. 221–233. Berkeley: University of California Press.

MIZON, G., and J.-F. RICHARD (1986): "The encompassing principle and its application to testing non-nested hypotheses," *Econometrica*, 54, 657–678.

NEWEY, W. (1985): "Maximum likelihood specification testing and conditional moment tests," *Econometrica*, 53, 1047–1070.

ROBINSON, P. M. (1983): "Nonparametric estimation for time series models," *Journal of Time Series Analysis*, 4, 185–208.

ROBINSON, P. M. (1991): "Consistent nonparametric entropy-based testing," *Review of Economic Studies*, 58, 437–453.

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

TAUCHEN, G. (1985): "Diagnostic testing and evaluation of maximum likelihood models," *Journal of Econometrics*, 30, 415–443.

White, H. (1982): "Maximum likelihood estimation of misspecified models," *Econometrica*, 50, 1–26.

White, H. (1987): "Specification testing in dynamic models," in *Advances in econometrics, Fifth World Congress*, ed. by T. F. Bewley, vol. 1, pp. 1–58. Cambridge University Press.

Wooldridge, J. (1992): "A test for functional form against nonparametric alternatives," *Econometric Theory*, 8, 452–475.

Zheng, J. (1996): "A consistent test of functional form via nonparametric estimation techniques," *Journal of Econometrics*, 75, 263–289.

A Appendix

A.1 Proof of Proposition 1:

We can write:

$$\begin{aligned}\sqrt{Th}\left(\tilde{\mathcal{G}}(Y) - \frac{1}{N}\sum_{s=1}^{N}\tilde{\mathcal{G}}(Y^s(\hat{\theta}))\right) &= \sqrt{Th}\left(\tilde{\mathcal{G}}(Y) - \mathcal{G}_0\right) \\ &- \frac{\sqrt{Th}}{N}\sum_{s=1}^{N}\left(\tilde{\mathcal{G}}(Y^s(\theta_0)) - \mathcal{G}_0\right) \\ &+ \frac{\sqrt{Th}}{N}\sum_{s=1}^{N}\left(\tilde{\mathcal{G}}(Y^s(\theta_0)) - \tilde{\mathcal{G}}(Y^s(\hat{\theta}))\right)\end{aligned}$$

The third term is $O_p(\sqrt{h})$ by virtue of (3). Under H_0, the first two terms are $O_p(1)$ by (4). By construction, the summands of the second term are independent among themselves and independent of the first term, so that

$$E\left|\sum_{s=1}^{N}\left(\tilde{\mathcal{G}}(Y^s(\theta_0)) - \mathcal{G}_0\right)\right|^2 = O(N^{-1}) = o(1),$$

and (4) suffices to conclude.

A.2 Proof of Theorem 1:

The proof follows as a corollary from Theorems 5.3 and 5.4 of Robinson (1983).

A.3 Proof of Theorem 2:

This is a generalization of a result stated in Berlinet, Gannoun, and Matzner-Lober (1998). Let $\mathcal{V}_i$ be a compact subset of $\mathbb{R}$ whose interior contains $Q(\zeta_i, p)$, and such that condition (ii) applies, and $f(Q(\zeta_i, p)|\zeta_i) > 0$ on $\mathcal{V}_i$. In view of Lemma 1, Lemma 2 and the continuous differentiability of $\hat{F}(.|\zeta_i)$, the Mean Value Theorem can be applied in the compact $\mathcal{V}_i$ between $\hat{Q}(\zeta_i, p)$ and $Q(\zeta_i, p)$, yielding:

$$\hat{Q}(\zeta_i, p) - Q(\zeta_i, p) = \frac{F(Q(\zeta_i, p)|\zeta_i) - \hat{F}(Q(\zeta_i, p)|\zeta_i)}{f(\tilde{Q}(\zeta_i, p)|\zeta_i)}$$

where $|\tilde{Q}(\zeta_i, p)|\zeta_i) - Q(\zeta_i, p)|\zeta_i)| \leq |\hat{Q}(\zeta_i, p)|\zeta_i) - Q(\zeta_i, p)|\zeta_i)|$. The result then follows from Lemmas 3-5.

<u>Lemma 1</u>:

Under the conditions of the above theorem, $\hat{Q}(\zeta_i, p)$ converges in probability to $Q(\zeta_i, p)$ for all $i = 1, \dots, d$.

<u>Proof of Lemma 1</u>:

The result follows from the fact that, for all $i = 1, \dots, d$ $F(.|\zeta_i)$ admits a unique quantile at level p, so that for all $\varepsilon > 0$, there exists some $\eta > 0$ such that

$$\begin{aligned} P(|\hat{Q}(\zeta_i, p) - Q(\zeta_i, p)| > \varepsilon) &\leq P(|F(\hat{Q}(\zeta_i, p)|\zeta_i) - F(Q(\zeta_i, p)|\zeta_i)| > \eta) \\ &\leq P(\sup_{y \in \mathbb{R}} |\hat{F}(y|\zeta_i) - F(y|\zeta_i)| > \eta) \end{aligned}$$

The result follows from Lemma 2 below.

<u>Lemma 2</u>

Let ξ be in $\mathcal{V}_i$.

$\hat{F}(\xi|\zeta_i)$ (resp. $\hat{f}(\xi|\zeta_i)$) converges in probability to $F(\xi|\zeta_i)$, (resp. $f(\xi|\zeta_i)$), uniformly on $\mathcal{V}_i$.

<u>Proof of Lemma 2</u>

$$\begin{aligned} |\hat{F}(\xi|\zeta_i) - F(\xi|\zeta_i)| &\leq \Big\{ |\hat{\phi}(\xi, \zeta_i) - E\hat{\phi}(\xi, \zeta_i)| + |E\hat{\phi}(\xi, \zeta_i) - \phi(\xi, \zeta_i)| \\ &+ |[1; \zeta_i] - E[1; \zeta_i]| + |E[1; \zeta_i] - f(\zeta_i)| \Big\} / [1; \zeta_i] \end{aligned}$$

The bias terms are $o(1)$ from the proof of lemma 3, and the third term is $o_p(1)$ as an immediate corollary to Theorem 4.1 of Robinson (1983).

As to the first term (denoted Φ), call

$$a_t = \int_{-\infty}^{\xi} (Th|h^{(i)}|)^{-1} l(h^{-1}(u-Y_t)) K_i(\zeta_i - Z_t; h^{(i)})\, du$$
$$- E\int_{-\infty}^{\xi} (Th|h^{(i)}|)^{-1} l(h^{-1}(u-Y_t)) K_i(\zeta_i - Z_t; h^{(i)})\, du.$$

For some $p > 2$, we have:

$$E|a_t|^p \leq 2E\left|\int_{-\infty}^{\xi} (Th|h^{(i)}|)^{-1} l(h^{-1}(u-Y_t)) K_i(\zeta_i - Z_t; h^{(i)})|\, du\right|^p$$
$$= O(T^{-p}|h^{(i)}|^{1-p}).$$

Therefore,

$$E\Phi^2 = E(\sum_{t=1+\tau_n}^{T} a_t)^2 = O\left(T(T^{-p}|h^{(i)}|^{1-p})\sum_{t=1}^{T} \alpha_t^{1/r}\right) = o(1).$$

using Davydov's inequality with $2/p = 1 - 1/r$ and Assumption 3.1.

The convergence in probability of $\hat{f}(\xi|\zeta_i)$ to $f(\xi|\zeta_i)$ is a corollary of Theorem 6.1 of Robinson (1983).

By construction, $\hat{F}(\xi|\zeta_i)$ is continuously differentiable with respect to ξ, so that both $\hat{F}$ and $\hat{f}$ are continuous on the compact $\mathcal{V}_i$ and therefore uniformly continuous. The result follows.

Lemma 3

$F(Q(\zeta_i,p)|\zeta_i) - E(\hat{F}(Q(\zeta_i,p)|\zeta_i)) = o((T|h^{(i)}|)^{-1/2})$, all $i = 1, \ldots, d$.

Proof of Lemma 3

Take ζ_i and $\xi \in \mathcal{V}_i$.

We have:

$$E\hat{\phi}(\xi,\zeta_i) = \int_{\mathbb{R}^{n+1}} \int_{-\infty}^{\xi} (h|h^{(i)}|)^{-1} l(h^{-1}(u-\omega)) K_i(\zeta_i - \lambda; h^{(i)})\ f(\omega,\lambda)\ du\, d\lambda\, d\omega$$
$$= \int_{\mathbb{R}^{n+1}} \int_{-\infty}^{\xi} l(\omega) K_i(\lambda;1) f(u - h\omega, \zeta_i - (h^{(i)}\lambda))\, du\, d\lambda\, d\omega$$
$$= \phi(\xi,\zeta_i) + \int_{-\infty}^{\xi} \left\{ \frac{h^2}{2} f^{(\xi\xi)}(u,\zeta_i) \int_{\mathbb{R}} \omega^2 l(\omega)\, d\omega \right.$$
$$+ \left. \sum_{j=1}^{n} \frac{h_{ij}^2}{2} f_{jj}^{(\zeta\zeta)}(u,\zeta_i) \int_{\mathbb{R}} \omega^2 k_{ij}(\omega)\, d\omega \right\} du + o(h^2 + \max_j(h_{ij}^2))$$
$$= \phi(\xi,\zeta_i) + O(h^2 + \max_j(h_{ij}^2)).$$

In the same way,

$$E[1;\zeta_i] = f(\zeta_i) + \sum_{j=1}^{n} \frac{h_{ij}^2}{2} f_{jj}^{(\zeta\zeta)}(\zeta_i) \int_{\mathbb{R}} \omega^2 k_{ij}(\omega)\, d\omega + o(\max_j(h_{ij}^2)).$$

Now,

$$\begin{aligned} F(\xi|\zeta_i) - E(\hat{F}(\xi|\zeta_i)) &= \frac{\phi(\xi,\zeta_i) - E\hat{\phi}(\xi,\zeta_i)}{f(\zeta_i)} - (f(\zeta_i) - E[1;\zeta_i])\frac{\phi(\xi,\zeta_i)}{f^2(\zeta_i)} \\ &+ O((\phi(\xi,\zeta_i) - E\hat{\phi}(\xi,\zeta_i))(f(\zeta_i) - E[1;\zeta_i])) \end{aligned}$$

Therefore,

$$F(\xi|\zeta_i) - E(\hat{F}(\xi|\zeta_i)) = O(h^2 + \max(h_{ij}^2)) = O(||h^{(i)}||^2) = o((T|h^{(i)}|)^{-1/2})$$

which proves the result.

Lemma 4

The d dimensional vector with elements

$$\left(\frac{T|h^{(i)}|[1;\zeta_i]}{p(1-p)\int \Pi_{j=1}^n k_{ij}^2(u)du_j}\right)^{1/2} \left\{F(Q(\zeta_i,p)|\zeta_i) - \hat{F}(Q(\zeta_i,p)|\zeta_i)\right\}$$

converges in distribution to a vector of independent standard normal variables.

Proof of Lemma 4

Using a first order expansion of the linear operator $(x_1, x_2) \to x_2/x_1$ and Lemma 5, we can write

$$\begin{aligned} \frac{\hat{\phi}(\xi,\zeta_i)}{[1;\zeta_i]} - \frac{E\hat{\phi}(\xi,\zeta_i)}{E[1;\zeta_i]} &= \left(-\frac{E\hat{\phi}(\xi,\zeta_i)}{(E[1;\zeta_i])^2} \quad \frac{1}{E[1;\zeta_i]}\right) \begin{pmatrix} [1;\zeta_i] - E[1;\zeta_i] \\ \\ \hat{\phi}(\xi,\zeta_i) - E\hat{\phi}(\xi,\zeta_i) \end{pmatrix} \\ &+ o_p\left((T|h^{(i)}|)^{-\frac{1}{2}}\right) \end{aligned}$$

From Lemma 5 below, the first term on the right hand side, normalized by

$$\left(\frac{T|h^{(i)}|}{\int \Pi_{j=1}^n k_{ij}^2(u)du_j}\right)^{1/2},$$

converges to a normal random variable with variance

$$\left(-\frac{\phi(\xi,\zeta_i)}{f(\zeta_i)^2} \quad \frac{1}{f(\zeta_i)}\right) \begin{pmatrix} f(\zeta_i) & \phi(\xi,\zeta_i) \\ \\ \phi(\xi,\zeta_i) & \phi(\xi,\zeta_i) \end{pmatrix} \begin{pmatrix} -\frac{\phi(\xi,\zeta_i)}{f(\zeta_i)^2} \\ \\ \frac{1}{f(\zeta_i)} \end{pmatrix}$$

which, taking ξ equal to $Q(\zeta_i, p)$, is $p(1-p)/f(\zeta_i)$.

The result follows from the observation that

$$\frac{\hat{\phi}(\xi,\zeta_i)}{[1;\zeta_i]} = \frac{\hat{\phi}(\xi,\zeta_i)}{E[1;\zeta_i]}\left(\frac{1}{1-\frac{E[1;\zeta_i]-[1;\zeta_i]}{E[1;\zeta_i]}}\right) = \frac{\hat{\phi}(\xi,\zeta_i)}{E[1;\zeta_i]} + o_p\left((T|h^{(i)}|)^{-\frac{1}{2}}\right)$$

Taking expectations, this allows to replace $E\frac{\hat{\phi}(\xi,\zeta_i)}{[1;\zeta_i]}$ by $\frac{E\hat{\phi}(\xi,\zeta_i)}{E[1;\zeta_i]}$ above.

Lemma 5

$$\left(\frac{T|h^{(i)}|}{\int \Pi_{j=1}^n k_{ij}^2(u)du_j}\right)^{1/2}\begin{pmatrix}[1;\zeta_i]-E[1;\zeta_i]\\ \\ \hat{\phi}(\xi,\zeta_i)-E\hat{\phi}(\xi,\zeta_i)\end{pmatrix} \to N\left(0,\begin{pmatrix} f(\zeta_i) & \phi(\xi,\zeta_i)\\ \\ \phi(\xi,\zeta_i) & \phi(\xi,\zeta_i)\end{pmatrix}\right)$$

Proof of Lemma 5

Let λ_1 and λ_2 be two real numbers. Call

$$\begin{aligned} V_{it} &= \lambda_1\left[K_i(\zeta_i - Z_t; h^{(i)}) - EK_i(\zeta_i - Z_t; h^{(i)})\right] \\ &+ \lambda_2\left[K_i(\zeta_i - Z_t; h^{(i)})\int_{-\infty}^{\xi} l(h^{-1}(u-Y_t))du \right.\\ &- \left. EK_i(\zeta_i - Z_t; h^{(i)})\int_{-\infty}^{\xi} l(h^{-1}(u-Y_t))du\right] \end{aligned}$$

We will prove the following propositions:

$$E|V_{it}V_{i,t+s}| \le C|h^{(i)}|^2, \quad \text{all } s>0; \tag{10}$$

for some fixed positive constant C, and

$$\lim_{T\to\infty}\frac{E|V_{it}|^2}{|h^{(i)}|} = \left(\lambda_2^2\phi(\xi,\zeta_i) + \lambda_1^2 f(\zeta_i) + 2\lambda_1\lambda_2\phi(\xi,\zeta_i)\right)\int \Pi_{j=1}^n k_{ij}^2(u)du_j \tag{11}$$

Because the kernels are bounded, the result follows from Lemma 7.1 of Robinson (1983).

For the proof of (11), let's consider for example the term in λ_2^2:

$$\begin{aligned} & |h^{(i)}|^{-1}E|K_i(\zeta_i - Z_t; h^{(i)})\int_{-\infty}^{\xi} l(h^{-1}(u-Y_t))du|^2 \\ =\ & \int_{\mathbb{R}^{n+1}} |h^{(i)}|^{-1}\left(K_i(\zeta_i-\lambda; h^{(i)})\int_{-\infty}^{\xi} l(h^{-1}(u-\omega))du\right)^2 f(\omega,\lambda)\,d\lambda\,d\omega \\ =\ & \int_{\mathbb{R}^n}(K_i(\lambda))^2\,d\lambda \int\left(\int_{-\infty}^{\xi} l(h^{-1}(u-\omega))du\right)^2 f(\omega,\zeta_i)\,d\omega + O(|h^{(i)}|) \\ \to\ & \int \Pi_{j=1}^n k_{ij}^2(u)du_j \int \mathbb{I}_{\{\omega\le\xi\}}\ f(\omega,\zeta_i)\,d\omega = \phi(\xi,\zeta_i)\int \Pi_{j=1}^n k_{ij}^2(u)du_j; \end{aligned}$$

and the remaining terms are treated in the same way. Finally, Lemma 8.3 of Robinson (1983) is used for the proof of (10).

A.4 Proof of Theorem 3:

The proof follows as a corollary from Theorem 2.