Nonparametric Bootstrap Tests for Neglected Nonlinearity in Time Series Regression Models*

Tae-Hwy Lee
Department of Economics
University of California
Riverside, CA 92521
taelee@mail.ucr.edu

Aman Ullah
Department of Economics
University of California
Riverside, CA 92521
ullah@mail.ucr.edu

June 2000

Abstract

Various nonparametric kernel regression estimators are presented, based on which we consider two nonparametric tests for neglected nonlinearity in time series regression models. One of them is the goodness-of-fit test of Cai, Fan, and Yao (2000) and another is the nonparametric conditional moment test by Li and Wang (1998) and Zheng (1996). Bootstrap procedures are used for these tests and their performance is examined via monte carlo experiments, especially with conditionally heteroskedastic errors.

Key Words: nonparametric test, nonlinearity, time series, functional-coefficient model, conditional moment test, naive bootstrap, wild bootstrap, conditional heteroskedasticity, GARCH, monte carlo.

*We would like to thank Zongwu Cai, Jianqing Fan, Qi Li, an anonymous referee, and Associate Editor for their helpful comments which greatly improved the paper and/or for sharing their programs. All errors are our own. Lee thanks UC Regents's Faculty Fellowship and Faculty Development Awards, and Ullah thanks the Academic Senate of UCR for the research support.

1 Introduction

Much research in empirical and theoretical econometrics has been centered around the estimation and testing of various functions such as regression functions (e.g., conditional mean and variance) and density functions. A traditional approach to studying these functions has been to first impose a parametric functional form and then proceed with the estimation and testing of interest. A major disadvantage of this approach is that the econometric analysis may not be robust to the slight data inconsistency with the particular parametric specification and this may lead to erroneous conclusions. In view of these problems, in the last four decades or so a vast amount of literature has appeared on the nonparametric and semiparametric approaches to econometrics, e.g., see the books by Härdle (1990), Fan and Gijbels (1996), and Pagan and Ullah (1999). The basic point in the nonparametric approach to econometrics is to realize that, in many instances, one is attempting to estimate an expectation of one variable, y, conditional upon others, x. This identification directs attention to the need to be able to estimate the conditional mean of y given x from the data y_t and x_t, $t = 1, \ldots, n$. A nonparametric estimate of this conditional mean simply follows as a weighted average $\sum_t w(x_t, x)y_t$, where $w(x_t, x)$ are a set of weights that depend upon the distance of x_t from the point x at which the conditional expectation is to be evaluated.

Based on these nonparametric estimation techniques of the conditional expectations, in recent years a rich literature has evolved on the consistent model specification tests in econometrics. For example, various test statistics for testing a parametric functional form have been proposed by Bierens (1982), Ullah (1985), Robinson (1989), Eubank and Spiegelman (1990), Yatchew (1992), Wooldridge (1992), Gozalo (1993), Härdle and Mammen (1993), Hong and White (1995), Zheng (1996), Bierens and Ploberger (1997), and Li and Wang (1998). Also, see Ullah and Vinod (1993), Whang and Andrews (1993), Delgado and Stengos (1994), Lewbel (1993, 1995), Aït-Sahalia et al (1994), Fan and Li (1996), Lavergne and Vuong (1996), and Linton and Gozalo (1997) for testing problems related to insignificance of regressors, non-nested hypothesis, semiparametric versus nonparametric regression models, among others. Most of these tests, especially the test for a parametric specification, are developed under the following goodness of fit measures: (i) compare the expected values of the squared error under the null and alternative hypotheses (e.g., Ullah (1985) type F statistic), (ii) calculate the expected value of the squared distance between the null and alternative model specifications (e.g., Härdle and Mammen (1993), Ullah and Vinod (1993), Aït-Sahalia (1994)), and (iii) calculate the expected value of the product of the error under the null with the model specified under the alternative (e.g., conditional moment tests of Bierens (1982), Zheng (1996), Fan and Li (1996), and Li and Wang (1998). All these three alternative goodness of fit measures are equal to zero under the null hypothesis of correct specification. For details, see Pagan and Ullah (1999).

We note here that the asymptotic as well as the simulation based finite sample properties of the most of the above mentioned test statistics have been extensively analyzed for the cross sectional models with independent data. However, not much is known about the asymptotic as well as the small sample performance of these test statistics for the case of time series models with weak dependent data, although see the recent works of Chen and Fan (1999), Hjellvik and Tjøstheim (1995, 1998), Hjellvik et al (1999), Kreiss et al (1998), Berg and Li (1998) and a very important contribution by Li (1999) where he develops the asymptotic theory results of Li-Wang-Zheng (LWZ) test under the goodness of fit measure (iii). The modest goal of this paper is to conduct an extensive monte carlo study to analyze the size and power properties of two kernel based tests for time series models. One of them is the bootstrap version of Ullah-type goodness of fit test (i) due to Cai, Fan, and Yao (2000, henceforth CFY), and another is the nonparametric conditional moment goodness of fit test (iii) of LWZ. We examine the bootstrap performances of these two goodness of fit tests because of the asymptotic validity results of using bootstrap methods for these statistics due to CFY (2000) and Berg and Li (1998). Berg and Li (1998) also support the better performance of LWZ over the Härdle and Mammen (1993) type tests considered for time series data in Hjellvik and Tjøstheim (1995, 1998), Hjellvik et al (1998), and Kreiss et al (1998). For the purpose of our simulation study we consider the testing of linearity against a large class of nonlinear time series models which include threshold autoregressive, bilinear, exponential autoregressive models, smooth transition autoregressive models, GARCH models, and various nonlinear autoregressive and moving average models. Both naive bootstrap and wild bootstrap procedures are used for our analysis. We also compare the bootstrap results with the results using the asymptotic distribution for LWZ test.

The plan of the paper is as follows. In Section 2, we present the nonparametric kernel regression estimators and the tests of CFY and LWZ based on them. Then in Section 3, we present the monte carlo results. Finally, Section 4 gives conclusions.

2 Nonparametric regression and specification testing

2.1 Nonparametric regression

Let $\{y_t, x_t\}, t = 1, \ldots, n$, be stochastic processes, where y_t is a scalar and $x_t = (x_{t1}, \ldots, x_{tk})$ is a $1 \times k$ vector which may contain the lagged values of y_t. Consider the regression model

$$y_t = m(x_t) + u_t \tag{1}$$

where $m(x_t) = E(y_t|x_t)$ is the true but unknown regression function and u_t is the error term such that $E(u_t|x_t) = 0$.

If $m(x_t) = g(x_t, \delta)$ is a correctly specified family of parametric regression functions then $y_t = g(x_t, \delta) + u_t$

is a correct model and, in this case, one can construct a consistent least squares (LS) estimator of $m(x_t)$ given by $g(x_t, \hat{\delta})$, where $\hat{\delta}$ is the LS estimator of the parameter δ.

In general, if the parametric regression $g(x_t, \delta)$ is incorrect or the form of $m(x_t)$ is unknown then $g(x_t, \hat{\delta})$ may not be a consistent estimator of $m(x_t)$. For this case, an alternative approach to estimate the unknown $m(x_t)$ is to use the consistent nonparametric kernel regression estimator which is essentially a local constant LS (LCLS) estimator. To obtain this estimator take Taylor series expansion of $m(x_t)$ around x so that

$$\begin{aligned} y_t &= m(x_t) + u_t \\ &= m(x) + e_t \end{aligned} \tag{2}$$

where $e_t = (x_t - x)m^{(1)}(x) + \frac{1}{2}(x_t - x)^2 m^{(2)}(x) + \cdots + u_t$ and $m^{(s)}(x)$ represents the s-th derivative of $m(x)$ at $x_t = x$. The LCLS estimator can then be derived by minimizing

$$\sum_{t=1}^{n} e_t^2 K_{tx} = \sum_{t=1}^{n} (y_t - m(x))^2 K_{tx} \tag{3}$$

with respect to constant $m(x)$, where $K_{tx} = K\left(\frac{x_t - x}{h}\right)$ is a decreasing function of the distances of the regressor vector x_t from the point $x = (x_1, \ldots, x_k)$, and $h \to 0$ as $n \to \infty$ is the window width (smoothing parameter) which determines how rapidly the weights decrease as the distance of x_t from x increases. The LCLS estimator so estimated is

$$\hat{m}(x) = \frac{\sum_{t=1}^{n} y_t K_{tx}}{\sum_{t=1}^{n} K_{tx}} = (\mathbf{i}'\mathbf{K}(x)\mathbf{i})^{-1}\mathbf{i}'\mathbf{K}(x)\mathbf{y} \tag{4}$$

where $\mathbf{K}(x)$ is the $n \times n$ diagonal matrix with the diagonal elements K_{tx} $(t = 1, \ldots, n)$, $\mathbf{i}$ is an $n \times 1$ column vector of unit elements, and $\mathbf{y}$ is an $n \times 1$ vector with elements y_t $(t = 1, \ldots, n)$. The estimator $\hat{m}(x)$ is due to Nadaraya (1964) and Watson (1964) (NW) who derived this in an alternative way. Generally $\hat{m}(x)$ is calculated at the data points x_t, in which case we can write the leave-one out estimator as

$$\hat{m}(x) = \frac{\sum_{t'=1, t' \neq t}^{n} y_{t'} K_{t't}}{\sum_{t'=1, t' \neq t}^{n} K_{t't}}, \tag{5}$$

where $K_{t't} = K\left(\frac{x_{t'} - x_t}{h}\right)$. The assumption that $h \to 0$ as $n \to \infty$ gives $x_t - x = O(h) \to 0$ and hence $Ee_t \to 0$ as $n \to \infty$. Thus the estimator $\hat{m}(x)$ will be consistent under certain smoothing conditions on h, K, and $m(x)$. In small samples however $Ee_t \neq 0$ so $\hat{m}(x)$ will be a biased estimator, see Pagan and Ullah (1999) for details on asymptotic and small sample properties.

An estimator which has a better small sample bias and hence the mean square error (MSE) behavior is the local linear LS (LLLS) estimator due to Stone (1977) and Cleveland (1979), also see Fan and Gijbels (1996) and Ruppert and Wand (1994) for their properties. In the LLLS estimator we take first order Taylor-Series expansion of $m(x_t)$ around x so that

$$y_t = m(x_t) + u_t = m(x) + (x_t - x)m^{(1)}(x) + v_t \tag{6}$$

$$
\begin{aligned}
&= \alpha(x) + x_t\beta(x) + v_t \\
&= X_t\delta(x) + v_t
\end{aligned}
$$

where $X_t = (1 \ x_t)$ and $\delta(x) = [\alpha(x) \ \beta(x)']'$ with $\alpha(x) = m(x) - x\beta(x)$ and $\beta(x) = m^{(1)}(x)$. The LLLS estimator of $\delta(x)$ is then obtained by minimizing

$$
\sum_{t=1}^{n} v_t^2 K_{tx} = \sum_{t=1}^{n} (y_t - X_t\delta(x))^2 K_{tx} \tag{7}
$$

and it is given by

$$
\tilde{\delta}(x) = (\mathbf{X}'\mathbf{K}(x)\mathbf{X})^{-1}\mathbf{X}'\mathbf{K}(x)\mathbf{y}. \tag{8}
$$

where $\mathbf{X}$ is an $n \times (k+1)$ matrix with the tth row X_t $(t = 1, \ldots, n)$.

The LLLS estimator of $\alpha(x)$ and $\beta(x)$ can be calculated as $\tilde{\alpha}(x) = (1 \ 0)\tilde{\delta}(x)$ and $\tilde{\beta}(x) = (0 \ 1)\tilde{\delta}(x)$. This gives

$$
\tilde{m}(x) = (1 \ x)\tilde{\delta}(x) = \tilde{\alpha}(x) + x\tilde{\beta}(x). \tag{9}
$$

Obviously when $X = \mathbf{i}$, $\tilde{\delta}(x)$ reduces to the NW's LCLS estimator $\hat{m}(x)$. An extension of the LLLS is the local polynomial LS (LPLS) estimators, see Fan and Gijbels (1996).

In fact one can obtain the local estimators of a general nonlinear model $g(x_t, \delta)$ by minimizing

$$
\sum_{t=1}^{n} [y_t - g(x_t, \delta(x))]^2 K_{tx} \tag{10}
$$

with respect to $\delta(x)$. For $g(x_t, \delta(x)) = X_t\delta(x)$ we get the LLLS in (8). Further when $h = \infty, K_{tx} = K(0)$ is a constant so that the minimization of $K(0)\sum[y_t - g(x_t, \delta(x))]^2$ is the same as the minimization of $\sum[y_t - g(x_t, \delta)]^2$, that is the local LS becomes the global LS estimator $\hat{\delta}$.

The LLLS estimator in (8) can also be interpreted as the estimator of the functional coefficient (varying coefficient) linear regression model

$$
\begin{aligned}
y_t &= m(x_t) + u_t \qquad (11) \\
&= X_t\delta(x_t) + u_t
\end{aligned}
$$

where $\delta(x_t)$ is approximated locally by a constant $\delta(x_t) \simeq \delta(x)$. The minimization of $\sum u_t^2 K_{tx}$ with respect to $\delta(x)$ then gives the LLLS estimator in (8), which can be interpreted as the LC varying coefficient estimator. An extension of this is to consider the linear approximation $\delta(x_t) \simeq \delta(x) + D(x)(x_t - x)'$ where $D(x) = \frac{\partial \delta(x_t)}{\partial x_t'}$ evaluated at $x_t = x$. In this case

$$
\begin{aligned}
y_t &= m(x_t) + u_t = X_t\delta(x_t) + u_t \qquad (12) \\
&\simeq X_t\delta(x) + X_t D(x)(x_t - x)' + u_t
\end{aligned}
$$

$$= X_t\delta(x) + [(x_t - x)\otimes X_t]vecD(x) + u_t$$
$$= X_t^x\delta^x(x) + u_t$$

where $X_t^x = [X_t \;\; (x_t - x)\otimes X_t]$ and $\delta^x(x) = [\delta(x)' \;\; (vecD(x))']'$. The LL varying coefficient estimator of $\delta^x(x)$ can then be obtained by minimizing

$$\sum_{t=1}^{n}[y_t - X_t^x\delta^x(x)]^2 K_{tx} \tag{13}$$

with respect to $\delta^x(x)$ as

$$\dot{\delta}^x(x) = (\mathbf{X}^{x\prime}\mathbf{K}(x)\mathbf{X}^x)^{-1}\mathbf{X}^{x\prime}\mathbf{K}(x)\mathbf{y}. \tag{14}$$

From this $\dot{\delta}(x) = (\mathbf{I}\; 0)\dot{\delta}^x(x)$, and hence

$$\dot{m}(x) = (1\; x\; 0)\dot{\delta}^x(x) = (1\; x)\dot{\delta}(x). \tag{15}$$

The above idea can be extended to the situations where $\xi_t = (x_t\; z_t)$ such that

$$E(y_t|\xi_t) = m(\xi_t) = m(x_t, z_t) = X_t\delta(z_t), \tag{16}$$

where the coefficients are varying with respect to only a subset of ξ_t; z_t is $1\times l$ and ξ_t is $1\times p$, $p = k + l$. Examples of these include functional coefficient autoregressive model (Chen and Tsay 1993, CFY 2000), smooth coefficient model (Li, Huang, and Fu 1997), random coefficient model (Raj and Ullah 1981), smooth transition autoregressive model (Granger and Teräsvirta 1993), exponential autoregressive model (Haggan and Ozaki 1981), and threshold autoregressive model (Tong 1990). Also see Section 3.

To estimate $\delta(z_t)$ we can again do a local constant approximation $\delta(z_t) \simeq \delta(z)$ and then minimize $\sum[y_t - X_t\delta(z)]^2 K_{tz}$ with respect to $\delta(z)$, where $K_{tz} = K(\frac{z_t - z}{h})$. This gives the LC varying coefficient estimator

$$\tilde{\delta}(z) = (\mathbf{X}'\mathbf{K}(z)\mathbf{X})^{-1}\mathbf{X}'\mathbf{K}(z)\mathbf{y} \tag{17}$$

where $\mathbf{K}(z)$ is a diagonal matrix of $K_{tz}, t = 1,\ldots,n$. When $z = x$, (17) reduces to the LLLS estimator $\tilde{\delta}(x)$ in (8).

CFY (2000) consider a local linear approximation $\delta(z_t) \simeq \delta(z) + D(z)(z_t - z)'$. The LL varying coefficient estimator of CFY is then obtained by minimizing

$$\begin{aligned}\sum_{t=1}^{n}[y_t - X_t\delta(z_t)]^2 K_{tz} &= \sum_{t=1}^{n}[y_t - X_t\delta(z) - [(z_t - z)\otimes X_t]vecD(z)]^2 K_{tz} \\ &= \sum_{t=1}^{n}[y_t - X_t^z\delta^z(z)]^2 K_{tz}\end{aligned} \tag{18}$$

with respect to $\delta^z(z) = [\delta(z)' \;\; (vecD(z))']'$ where $X_t^z = [X_t \;\; (z_t - z)\otimes X_t]$. This gives

$$\ddot{\delta}^z(z) = (\mathbf{X}^{z\prime}\mathbf{K}(z)\mathbf{X}^z)^{-1}\mathbf{X}^{z\prime}\mathbf{K}(z)\mathbf{y}, \tag{19}$$

and $\ddot{\delta}(z) = (\mathbf{I}\ 0)\ddot{\delta}^z(z)$. Hence

$$\ddot{m}(\xi) = (1\ x\ 0)\ddot{\delta}^z(z) = (1\ x)\ddot{\delta}(z). \tag{20}$$

For the asymptotic properties of these varying coefficient estimators, see CFY (2000). When $z = x$, (19) reduces to the LL varying coefficient estimator $\dot{\delta}^x(x)$ in (14).

2.2 Nonparametric tests for functional forms

Consider the problem of testing a specified parametric model against a nonparametric (NP) alternative

$$\begin{aligned} H_o &: E(y_t|\xi_t) = g(\xi_t, \delta) \\ H_1 &: E(y_t|\xi_t) = m(\xi_t). \end{aligned}$$

In particular, if we are to test for neglected nonlinearity in the regression models, set $g(\xi_t, \delta) = \xi_t\delta$. Then under H_0, the process $\{y_t\}$ is linear in mean conditional on ξ_t

$$H_0 : P[E(y_t|\xi_t) = \xi_t\delta] = 1 \text{ for some } \delta \in \mathbb{R}^p. \tag{21}$$

The alternative of interest is the negation of the null, that is,

$$H_1 : P[E(y_t|\xi_t) = \xi_t\delta] < 1 \text{ for all } \delta \in \mathbb{R}^p. \tag{22}$$

When the alternative is true, a linear model is said to suffer from 'neglected nonlinearity'. Note that $\xi_t = (x_t\ z_t) = x_t$ when $z_t = x_t$.

Using the nonparametric estimation technique to construct consistent model specification tests was first suggested by Ullah (1985). The idea is to compare the parametric residual sum of squares (RSS$^{\mathrm{P}}$), $\sum \hat{u}_t^2$, $\hat{u}_t = y_t - g(\xi_t, \hat{\delta})$ with the nonparametric RSS (RSS$^{\mathrm{NP}}$), $\sum \tilde{u}_t^2$, where $\tilde{u}_t = y_t - \ddot{m}(\xi_t)$. The test statistic is

$$T = \frac{(\mathrm{RSS}^{\mathrm{P}} - \mathrm{RSS}^{\mathrm{NP}})}{\mathrm{RSS}^{\mathrm{NP}}} = \frac{\sum \hat{u}_t^2 - \sum \tilde{u}_t^2}{\sum \tilde{u}_t^2}, \tag{23}$$

or simply $T' = (\mathrm{RSS}^{\mathrm{P}} - \mathrm{RSS}^{\mathrm{NP}})$. We reject the null hypothesis when T is large. $\sqrt{n}T$ has a degenerate distribution under H_0. Yatchew (1992) avoids this degeneracy by splitting sample of n into n_1 and n_2 and calculating $\sum \hat{u}_t^2$ based on n_1 observations and $\sum \tilde{u}_t^2$ based on n_2 observations. Lee (1992) uses density weighted residuals and compares $\sum w_t \hat{u}_t^2$ with $\sum \tilde{u}_t^2$. Fan and Li (1992) uses different normalizing factor and show the asymptotic normality of $nh^{p/2}T'$.

An alternative way is to use the bootstrap method as suggested by CFY (2000). The bootstrap allows the implementation of (23) and it involves the following steps to evaluate p-values of T to test for $g(\xi_t, \delta) = X_t\delta$.

1. Generate the bootstrap residuals $\{\tilde{u}_t^*\}$ from the centered NP residuals $(\tilde{u}_t - \bar{u})$ where $\bar{u} = n^{-1}\sum \tilde{u}_t$.

(a) For naive bootstrap, $\{\tilde{u}_t^*\}$ is obtained from randomly resampling $\{\tilde{u}_t - \bar{u}\}$ with replacement.

(b) For wild bootstrap, $\tilde{u}_t^* = a(\tilde{u}_t - \bar{u})$ with probability $r = (\sqrt{5}+1)/2\sqrt{5}$ and $\tilde{u}_t^* = b(\tilde{u}_t - \bar{u})$ with probability $1-r$ $(t = 1, \ldots, n)$, where $a = -(\sqrt{5}-1)/2$ and $b = (\sqrt{5}+1)/2$. See Li and Wang (1998, pp. 150-151).

2. Generate the bootstrap sample $\{y_t^*\}_{t=1}^n$:

(a) When x_t is lagged dependent variables (e.g., see Blocks 1, 2, 5, 6 in Section 3), generate initial values of y_t^* for $t = 1, \ldots, k$, from $N(\bar{y}, \hat{\sigma}_Y^2)$, and then get $y_t^* \equiv X_t^* \hat{\delta} + \tilde{u}_t^*$ recursively for $t = k+1, \ldots, n$.

(b) When x_t is exogenous (Blocks 3, 4 in Section 3), then $x_t^* = x_t$ and $y_t^* \equiv X_t \hat{\delta} + \tilde{u}_t^*$ $(t = 1, \ldots, n)$.

3. Using the bootstrap sample $\{y_t^*\}_{t=1}^n$, calculate the bootstrap test statistic T^* using, for the sake of simplicity, the same h used in estimation with the original sample as done in CFY (2000).

4. Repeat the above steps B times and use the empirical distribution of T^* as the null distribution of T. We use $B = 500$. The bootstrap p-value of the test T is simply the relative frequency of the event $\{T^* \geq T\}$ in the bootstrap resamples.

Kreiss et al (1998) provide more detailed reasons why the bootstrap works in general nonparametric regression setting. They proved that asymptotically the conditional distribution of the bootstrap test statistic is indeed the distribution of the test statistic under the null hypothesis. As mentioned by CFY (2000) it may be proved that the similar result holds for T as long as $\hat{\delta}$ converges to δ at the rate $n^{-1/2}$. We use both naive bootstrap (Efron 1979) and wild bootstrap (Wu 1986, Liu 1988). The wild bootstrap method preserves the conditional heteroskedasticity in the original residuals. For wild bootstrap, see also Shao and Tu (1995, p. 292), Härdle (1990, p. 247), or Li and Wang (1998, p. 150).

Another test of the parametric specification follows from the combined regression

$$y_t = g(\xi_t, \delta) + E(u_t|\xi_t) + \varepsilon_t \tag{24}$$

where $\varepsilon_t = u_t - E(u_t|\xi_t)$ such that $E(\varepsilon_t|\xi_t) = 0$. The test for the parametric specification is then the conditional moment test for $E(u_t|\xi_t) = 0$, which is identical to testing

$$E[u_t E(u_t|\xi_t) f(\xi_t)] = 0, \tag{25}$$

where $f(\xi_t)$ is the density of ξ. A sample estimator of the left hand side of (25) is

$$\begin{aligned} L' &= \frac{1}{n} \sum_{t=1}^{n} \hat{u}_t E(\hat{u}_t|\xi_t) \hat{f}(\xi_t) \qquad (26) \\ &= \frac{1}{n(n-1)h^p} \sum_{t=1}^{n} \sum_{t'=1, t' \neq t}^{n} \hat{u}_t \hat{u}_{t'} K_{t't} \end{aligned}$$

where $E(\hat{u}_t|\xi_t) = \sum_{t' \neq t} \hat{u}_{t'} K_{t't} / \sum_{t' \neq t} K_{t't}$ from (5) and $\hat{f}(\xi_t) = [(n-1)h^p]^{-1} \sum_{t' \neq t} K_{t't}$ is the kernel density estimator; $K_{t't} = K(\frac{\xi_{t'} - \xi_t}{h})$. The asymptotic test statistic is then given by

$$L = nh^{p/2} \frac{L'}{\sqrt{\hat{\omega}}} \sim N(0, 1) \tag{27}$$

where $\hat{\omega} = 2(n(n-1)h^p)^{-1} \sum_t \sum_{t' \neq t} \hat{u}_t^2 \hat{u}_{t'}^2 K_{t't}^2$ is a consistent estimator of the asymptotic variance of $nh^{p/2} L'$, see Zheng (1996), Fan and Li (1996), Li and Wang (1998), Fan and Ullah (1999), and Rahman and Ullah (1999), for details. Also, see Pagan and Ullah (1999, Ch. 3) and Ullah (1999) for the relationship of this test statistic with other nonparametric specification tests. Based on the asymptotic results of Fan and Li (1996, 1997, 1999) and Li (1999) for dependent data, Berg and Li (1998) establish the asymptotic validity of using the wild bootstrap method for L for time-series. The bootstrap p-values for L to test for the adequacy of the linear parametric model, $g(\xi_t, \delta) = X_t \delta$, can be computed as follows.

1. Generate the bootstrap residuals $\{\hat{u}_t^*\}$ from $\hat{u}_t = y_t - X_t \hat{\delta}$:
   (a) For naive bootstrap, $\{\hat{u}_t^*\}$ is obtained from randomly resampling $\{\hat{u}_t\}$ with replacement.
   (b) For wild bootstrap, $\hat{u}_t^* = a\hat{u}_t$ with probability r and $\hat{u}_t^* = b\hat{u}_t$ with probability $1 - r$ as for T discussed above.
2. Generate the bootstrap sample $\{y_t^*\}_{t=1}^n$:
   (a) When x_t is lagged dependent variables (Blocks 1, 2, 5, 6), generate initial values of y_t^* for $t = 1, \ldots, k$, from $N(\bar{y}, \hat{\sigma}_Y^2)$, and then get $y_t^* \equiv X_t^* \hat{\delta} + \hat{u}_t^*$ recursively for $t = k+1, \ldots, n$.
   (b) When x_t is exogenous (Blocks 3, 4), then $x_t^* = x_t$ and $y_t^* \equiv X_t \hat{\delta} + \hat{u}_t^*$ $(t = 1, \ldots, n)$.
3. Using the bootstrap sample $\{y_t^*\}_{t=1}^n$, calculate the bootstrap test statistic L^*.
4. Repeat the above steps B times and use the empirical distribution of L^* as the null distribution of L. We use $B = 500$. The bootstrap p-value of the test L is the relative frequency of the event $\{L^* \geq L\}$ in the bootstrap resamples.

3 Monte carlo

In this section we examine the finite sample properties of the T test and the L test especially with the empirical null distributions being generated by the bootstrap method. Asymptotic critical values are also used for the L test. To generate data we use the following models, all of which have been used in the related literature. Most of them are univariate while there are some multivariate situations. There are six blocks.

The error term ε_t below is *i.i.d.* $N(0,1)$ unless otherwise is indicated. The models will be referred by the name in parentheses in bold.

BLOCK 1 (Lee, White, and Granger, 1993)

Linear (**AR**)

$$y_t = 0.6y_{t-1} + \varepsilon_t$$

Linear AR with GARCH (**AR′**)

$$\begin{aligned} y_t &= 0.6y_{t-1} + \varepsilon_t \\ h_t &\equiv E(\varepsilon_t^2 | y_{t-1}) = (1-\alpha-\beta) + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} \end{aligned}$$

Bilinear (**BL**)

$$y_t = 0.7y_{t-1}\varepsilon_{t-2} + \varepsilon_t$$

Threshold Autoregressive (**TAR**)

$$\begin{aligned} y_t &= 0.9y_{t-1} + \varepsilon_t \qquad |y_{t-1}| \leq 1 \\ &= -0.3y_{t-1} + \varepsilon_t \qquad |y_{t-1}| > 1 \end{aligned}$$

Sign Nonlinear Autoregressive (**SGN**)

$$y_t = \text{sign}(y_{t-1}) + \varepsilon_t$$

where $\text{sign}(x) = 1$ if $x > 0, 0$ if $x = 0$, and -1 if $x < 0$. This is a process examined in Granger and Teräsvirta (1999), which is a first-order nonlinear autoregressive model but has such misleading linear property that estimated autocorrelations are similar to those of a long-memory process.

Rational Nonlinear Autoregressive (**NAR**)

$$y_t = \frac{0.7\,|y_{t-1}|}{|y_{t-1}| + 2} + \varepsilon_t$$

BLOCK 2 (Lee, White, and Granger, 1993)

MA(2) (**M1**)

$$y_t = \varepsilon_t - 0.4\varepsilon_{t-1} + 0.3\varepsilon_{t-2}$$

Heteroskedastic MA(2) (**M2**)

$$y_t = \varepsilon_t - 0.4\varepsilon_{t-1} + 0.3\varepsilon_{t-2} + 0.5\varepsilon_t\varepsilon_{t-2}$$

Note that M2 is linear in conditional mean as the forecastable part of M2 is linear, and the final term introduces heteroskedasticity.

Nonlinear MA **(M3)**

$$y_t = \varepsilon_t - 0.3\varepsilon_{t-1} + 0.2\varepsilon_{t-2} + 0.4\varepsilon_{t-1}\varepsilon_{t-2} - 0.25\varepsilon_{t-2}^2$$

AR(2) **(M4)**

$$y_t = 0.4y_{t-1} - 0.3y_{t-2} + \varepsilon_t$$

Bilinear AR **(M5)**

$$y_t = 0.4y_{t-1} - 0.3y_{t-2} + 0.5y_{t-1}\varepsilon_{t-1} + \varepsilon_t$$

Bilinear ARMA **(M6)**

$$y_t = 0.4y_{t-1} - 0.3y_{t-2} + 0.5y_{t-1}\varepsilon_{t-1} + 0.8\varepsilon_{t-1} + \varepsilon_t$$

BLOCK 3 (Lee, White, and Granger, 1993)

Square (**SQ**)

$$y_t = x_t^2 + a_t$$

Exponential (**EXP**)

$$y_t = \exp(x_t) + a_t$$

These are bivariate models where $x_t = 0.6x_{t-1} + \varepsilon_t,\ a_t \sim N(0, 5^2)$,and a_t, ε_t are independent.

BLOCK 4 (Zheng, 1996)

Five models with $\xi_t = (x_{t1}\ x_{t2})$ are considered in this block. Let u_{t1} and u_{t2} be drawn from $IN(0,1)$. Two regressors x_{t1} and x_{t2} are defined as $x_{t1} = u_{t1}$ and $x_{t2} = (u_{t1} + u_{t2})/\sqrt{2}$.

Linear **(Z1)**

$$y_t = 1 + x_{t1} + x_{t2} + \varepsilon_t$$

Linear with conditionally heteroskedastic error **(Z1′)**

$$\begin{aligned} y_t &= 1 + x_{t1} + x_{t2} + \varepsilon_t \\ h_t &\equiv E(\varepsilon_t^2|\xi_t) = (1 + x_{t1}^2 + x_{t2}^2)/3 \end{aligned}$$

Quadratic **(Z2)**

$$y_t = 1 + x_{t1} + x_{t2} + x_{t1}x_{t2} + \varepsilon_t$$

Concave **(Z3)**

$$y_t = (1 + x_{t1} + x_{t2})^{1/3} + \varepsilon_t$$

Convex **(Z4)**

$$y_t = (1 + x_{t1} + x_{t2})^{5/3} + \varepsilon_t$$

BLOCK 5 (Cai, Fan, and Yao, 2000)

Exponential AR **(EXPAR)**

$$
\begin{aligned}
y_t &= a_1(y_{t-1})y_{t-1} + a_2(y_{t-1})y_{t-2} + \varepsilon_t \\
a_1(y_{t-1}) &= 0.138 + (0.316 + 0.982y_{t-1})\exp(-3.89y_{t-1}^2) \\
a_2(y_{t-1}) &= -0.437 - (0.659 + 1.260y_{t-1})\exp(-3.89y_{t-1}^2) \\
\varepsilon_t &\sim IN(0,\, 0.2^2)
\end{aligned}
$$

Threshold AR **(TAR)**

$$
\begin{aligned}
y_t &= a_1(y_{t-2})y_{t-1} + a_2(y_{t-2})y_{t-2} + \varepsilon_t \\
a_1(y_{t-2}) &= 0.4I(y_{t-2} \leq 1) - 0.8I(y_{t-2} > 6) \\
a_2(y_{t-2}) &= -0.6I(y_{t-2} \leq 1) + 0.2I(y_{t-2} > 1) \\
\varepsilon_t &\sim IN(0,\, 1)
\end{aligned}
$$

BLOCK 6 (Teräsvirta, Lin, and Granger, 1993)

Logistic smooth transition AR **(LSTAR)**

$$
\begin{aligned}
y_t &= 1.8y_{t-1} - 1.06y_{t-2} + (0.02 - 0.9y_{t-1} + 0.795y_{t-2})F(y_{t-1}) + \varepsilon_t \\
F(y_{t-1}) &= [1 + \exp\{-100(y_{t-1} - 0.02)\}]^{-1} \\
\varepsilon_t &\sim IN(0,\, 0.02^2)
\end{aligned}
$$

Exponential smooth transition AR **(ESTAR)**

$$
\begin{aligned}
y_t &= 1.8y_{t-1} - 1.06y_{t-2} + (-0.9y_{t-1} + 0.795y_{t-2})F(y_{t-1}) + \varepsilon_t \\
F(y_{t-1}) &= [1 - \exp\{-4000y_{t-1}^2\}]^{-1} \\
\varepsilon_t &\sim IN(0,\, 0.01^2)
\end{aligned}
$$

To estimate $\hat{u}_t$ for the linear model and $\tilde{u}_t$ for the NP model, the information set used are $\xi_t = y_{t-1}$ for Block 1, $\xi_t = (y_{t-1}\ y_{t-2})$ for Blocks 2, 5, and 6, $\xi_t = x_t$ for Block 3, and $\xi_t = (x_{t1}\ x_{t2})$ for Block 4.

For the T test, as suggested by CFY (2000), we select h using out-of-sample cross-validation. Let m and Q be two positive integers such that $n > mQ$. The basic idea is first to use Q sub-series of lengths $n - qm$ ($q = 1, \ldots, Q$) to estimate the coefficient functions $\delta_q(z_t)$ and then to compute the one-step forecast errors of the next segment of the time series of length m based on the estimated models. That is to choose h minimizing the average of the mean square forecast errors

$$AMS(h) = \sum_{q=1}^{Q} AMS_q(h) \tag{28}$$

where

$$AMS_q(h) = \frac{1}{m} \sum_{t=n-qm+1}^{n-qm+m} [y_t - X_t^z \ddot{\delta}_q^z(z)]^2 \tag{29}$$

and $\ddot{\delta}_q^z(\cdot)$ are computed from the sample $\{y_t\ \xi_t\}_{t=1}^{n-qm}$. We use $m = [0.1n]$, $Q = 4$, and the Epanechinikov kernel $K(z) = \frac{3}{4}(1 - z^2)1(|z| < 1)$. We use a scalar 'threshold variable' z_t (with $l = 1$) for all models: $z_t = y_{t-1}$ for Blocks 1, 2, and 6, $z_t = x_t$ for Block 3, and $z_t = x_{t1}$ for Block 4. For Block 5, $z_t = y_{t-1}$ for EXPAR and $z_t = y_{t-2}$ for TAR.

For the L test, as in Li and Wang (1998, p. 154), we use a standard normal kernel. Note that ξ_t is an $1 \times p$ vector, and $p = 1$ for Blocks 1, 3 and $p = 2$ for Blocks 2, 4, 5, 6. Thus the smoothing parameter h is chosen as $h_i = c\hat{\sigma}_i n^{-1/5}$ $(i = 1)$ for Blocks 1 and 3, and $h_i = c\hat{\sigma}_i n^{-1/6}$ $(i = 1, 2)$ for Blocks 2, 4, 5, 6, where $\hat{\sigma}_i$ is the sample standard deviation of i-th element of ξ. The three values of $c = 0.5$, 1, and 2 are used, and the corresponding estimated rejection probability will be denoted as L_c. In computing L, h^p shown in (26) and (27) is replaced with $\prod_{i=1}^{p} h_i$.

Test statistics are denoted as T^j and L_c^j, with the superscripts $j = A, B, W$ referring to the methods of obtaining the null distributions of the test statistics; asymptotics $(j = A)$, naive bootstrap $(j = B)$, and wild bootstrap $(j = W)$. Monte carlo experiments are conducted with 500 bootstrap resamples and 1000 monte carlo replications.

Table 1 gives the estimated size of the tests for the data generating processes (DGP) which are linear in conditional mean with the conditional homoskedastic errors. The size performance of the tests are different for dependent processes (AR, M1, M4) than for independent process (Z1). For Z1 process, the naive bootstrap CFY test T^B tends to under-reject the null while the wild bootstrap test T^W tends to over-reject the null. The LWZ tests work better for Z1 than for AR, M1, and M4, for all three values of c. For the three dependent processes (AR, M1, M4), both bootstrap procedures work relatively well with $c = 0.5$: $L_{0.5}$ is better than $L_{1.0}$ which is better than $L_{2.0}$, and the size of L is quite sensitive to the choice of c and hence bandwidth h. On the other hand, for the independent process Z1, the LWZ tests work well with all $c = 0.5, 1$, and 2. Both bootstrap tests L_c^B and L_c^W are better than the asymptotic test L_c^A. This tells that the optimal choice of c for time series is more important than for independent processes. The two bootstrap procedures are generally similar because the errors are homoskedastic in Table 1.

For parametric models, Davidson and MacKinnon (1999) show that the size distortion of a bootstrap test is at least of the order $n^{-1/2}$ smaller than that of the corresponding asymptotic test. For nonparametric models, h also enters the order of refinement. Li and Wang (1998) show that if the distribution of L^j ($j = A, B, W$) admit an Edgeworth expansion then the bootstrap distribution approximates the null distribution of L with an error of order $n^{-1/2}h^{p/2}$ improving over the normal approximation. Since L is asymptotically normal under the null, the bootstrap tests L^B and L^W are more accurate than the asymptotic test L^A, as confirmed in the simulation. See Hall (1992) for further discussion on Edgeworth expansions and the extent of the refinements in various contexts.

Table 2 gives the estimated size of the tests for the data generating processes (DGP) which are linear in conditional mean with conditional heteroskedastic errors. For AR$'$, we consider GARCH errors with five different parameter values: $(\alpha, \beta) = (0.5, 0.0)$, $(0.7, 0.0)$, $(0.1, 0.89)$, $(0.3, 0.69)$, and $(0.5, 0.49)$. The condition for the existence of the unconditional fourth moment is $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$ (Bollerslev, 1986). Accordingly, the condition is $\alpha < 0.577$ if $\beta = 0$; $\beta < 0.890$ if $\alpha = 0.1$; $\beta < 0.606$ if $\alpha = 0.3$; and $\beta < 0.207$ if $\alpha = 0.5$. Thus, for a given values of β or $\alpha + \beta$, the series becomes more leptokurtic as α increases. Table 2 shows that with $\beta = 0$ fixed, the size distortion is larger with the larger α. With $\alpha + \beta = 0.99$ fixed, the size distortion is larger also as α increases. The size distortion generally gets worse as n increases. This is most apparent with L^B as the naive bootstrap does not preserve the conditional heteroskedasticity in resampling.

Generally, as discussed in Lee et al (1993, p. 288), the conditional heteroskedasticity will have one of two effects: either it will cause the size of a test to be incorrect while still resulting in a test statistic bounded in probability under the null, or it will directly lead (asymptotically) to rejection despite linearity in mean. The test statistic L is a conditional moment test based on the fact that $E(u_t|\xi_t) = 0$ under the null hypothesis (21) which will then imply equation (25) for L. As this moment condition will hold even under the presence of the conditional heteroskedasticity (which can be shown by the law of iterated expectations), L should not have power to reject the null for the DGPs AR$'$ and Z1$'$ which are linear in conditional mean with conditionally heteroskedastic errors. Note that the asymptotic test L_c^A works well with the conditionally heteroskedastic errors. However, the size of the naive bootstrap test L_c^B is adversely affected by the conditional heteroskedasticity, which is more serious with a larger sample size.

Two remedies may be considered: one may either (1) remove the *effect* of the conditional heteroskedasticity or (2) remove the conditional heteroskedasticity *itself*. The first is relevant to L whose size is adversely affected. The effect of the conditional heteroskedasticity can be removed using a heteroskedasticity-consistent covariance matrix estimator or using the wild bootstrap that preserves the heteroskedasticity in resampling. We use the wild bootstrap here. The results in Table 2 show that the LWZ test with the wild bootstrap L_c^W generally has the adequate size for the both DGPs AR$'$ and Z1$'$.

On the other hand, T is not a conditional moment test as it is not based on any moment conditions. T is constructed to compare the two residual sums of squares $\mathrm{RSS}^{\mathrm{P}}$ and $\mathrm{RSS}^{\mathrm{NP}}$. As $\mathrm{RSS}^{\mathrm{NP}}$ is estimated from the functional coefficient (FC) model, if the FC model absorbs some of the conditional heteroskedasticity the size of the CFY test T will be incorrect, which we may observe in Table 2. Note that the size distortion generally tends to get more severe as n increases especially for AR$'$. The use of the wild bootstrap reduces the size distortion but only by small margin. In this case one may attempt the second remedy by removing the conditional heteroskedasticity *itself* whenever one is confidently able to specify the form of the conditional heteroskedasticity $h_t = var(y_t|\xi_t)$. Then we may compare the weighted parametric residual sum of squares ($\mathrm{WRSS}^{\mathrm{P}}$), $\sum \hat{u}_t^2/h_t$, $\hat{u}_t = y_t - g(\xi_t, \hat{\delta})$ with the weighted nonparametric RSS ($\mathrm{WRSS}^{\mathrm{NP}}$), $\sum \tilde{u}_t^2/h_t$, where $\tilde{u}_t = y_t - \ddot{m}(\xi_t)$. When h_t is a known function the CFY (2000) bootstrap procedure can be applied to the modified T statistic with the weighted RSS's. However, when h_t is unknown, it needs to be estimated. Use of misspecified conditional variance model in the procedure will again adversely affect the size of the test. Furthermore, if the alternative is true, the fitted conditional heteroskedasticity model can absorb some or even much of the neglected nonlinearity in conditional mean model. Conceivably, this could have adverse impact on the power of T statistic. Consideration of the second remedy together with the wild bootstrap could raise issues that take us well beyond the scope of the present study and their investigation is left for other work.

Table 3 presents the power of the tests T and L at 5% level. The results at 1% and 10% levels are available but not presented to save space. As the results obtained can be considerably influenced by the choice of nonlinear models, we try to include as many different types of nonlinear models as possible. Neither T nor L is uniformly superior to the other. T has relatively superior power for BL and ESTAR, and has power comparable to L in other cases.

4 Conclusions

We have presented a unified framework for various nonparametric kernel regression estimators, based on which we have considered two nonparametric tests T and L for neglected nonlinearity in regression models. We investigate them in several aspects: (1) T vs. L, (2) dependent process (AR) vs. independent process (Z1), (3) conditional homoskedasticity (AR and Z1) vs. conditional heteroskedasticity (AR$'$ and Z1$'$), (4) naive bootstrap (B) vs. wild bootstrap (W).

When the errors are conditionally heteroskedastic, the wild bootstrap LWZ test L^W works pretty well. However, the use of the wild bootstrap for T^W does not correct the size problem. This difference of the two statistics is due to the different construction of the test statistics: L is constructed based on a moment condition implying linearity in conditional *mean*, while T is constructed to detect any possible improvement in terms of residual variance via a nonparametric model over a linear model. Hence, the LWZ test can be robustified to the presence of conditional heteroskedasticy in testing for the linearity in conditional mean, while T will have power to detect neglected nonlinearity in conditional mean as well as the conditional heteroskedasticity. The choice of the bandwidth c in L_c is more important for time series processes than for independent process.

References

Aït-Sahalia, Y., Bickel, P.J., Stoker, T.M. (1994), "Goodness-of-Fit Tests for Regression Using Kernel Models", Princeton, UC Berkeley, and MIT, Submitted to *Econometrica.*

Berg, M.D., and D. Li (1998), "A Consistent Bootstrap Test for Time Series Regression Models", University of Guelph.

Bierens, H.J. (1982), "Consistent Model Specification Tests", *Journal of Econometrics*, 20, 105-134.

Bierens, H.J. and W. Ploberger (1997), "Asymptotic Theory of Integrated Conditional Moment Tests", *Econometrica*, 65, 1129-1151.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics,* 31, 307-327.

Cai, Z., J. Fan, and Q. Yao (2000), "Functional-coefficient Regression Models for Nonlinear Time Series", Forthcoming, *Journal of the American Statistical Association.*

Chen, X. and Y. Fan (1999), "Consistent Hypothesis Testing in Semiparametric and Nonparametric Models for Econometric Time Series", *Journal of Econometrics,* 91(2), 373-401.

Chen, R. and R.S. Tsay (1993), "Functional-coefficient Autoregressive Models", *Journal of American Statistical Association*, 88, 298-308.

Cleveland, W.S. (1979), "Robust Locally Weighted Regression and Smoothing Scatter Plots", *Journal of American Statistical Association,* 74, 829-836.

Davidson, R. and J.G. MacKinnon (1999), "The Size Distortion of Bootstrap Tests", *Econometric Theory,* 15(3), 361-376.

Delgado, M.A. and T. Stengos (1994), "Semiparametric Testing of Non-nested Econometric Models", *Review of Economic Studies*, 75, 345-367.

Efron, B. (1979), "Bootstrapping Methods: Another Look at the Jackknife", *Annals of Statistics*, 7, 1-26.

Eubank, R.L. and C.H. Spiegelman (1990), "Testing the Goodness-of-Fit of the Linear Models via Nonparametric Regression Techniques", *Journal of the American Statistical Association*, 85, 387-392.

Fan, J. and I. Gijbels (1996), *Local Polynomial Modelling and Its Applications*, Chapman and Hall, London.

Fan, Y. (1998), "Goodness of Fit Tests Based on Kernel Density Estimation with Fixed Smoothing Parameters", *Econometric Theory,* 14, 604-621.

Fan, Y. and Q. Li (1992), "Consistent Model Specification Tests: Omitted Variables and Semiparametric Functional Forms", Manuscript, University of Windsor.

Fan, Y. and Q. Li (1996), "Consistent Model Specification Tests: Omitted Variables and Semiparametric Functional Forms", *Econometrica,* 64, 865-890.

Fan, Y. and Q. Li (1997), "A Consistent Nonparametric Test for Linearity for AR(p) Models", *Economics Letters*, 55, 53-59.

Fan, Y. and Q. Li (1999), "Central Limit Theorem for Degenerate U-Statistics of Absolutely Regular Processes with Applications to Model Specification Tests", *Journal of Nonparametric Statistics*, 10, 245-271.

Fan, Y. and A. Ullah (1999), "Asymptotic Normality of a Combined Regression Estimator", forthcoming, *Journal of Multivariate Analysis.*

Gozalo, P.L. (1993), "A Consistent Model Specification Test for Nonparametric Estimation of Regression Function Models", *Econometric Theory*, 9, 451-477.

Granger, C.W.J. and T. Teräsvirta (1993), *Modelling Nonlinear Economic Relationships,* Oxford University Press, New York.

Granger, C.W.J. and T. Teräsvirta (1999), "A Simple Nonlinear Time Series Model with Misleading Linear Properties", *Economics Letters*, 62, 161-165.

Haggan, V. and T. Ozaki (1981), "Modeling Nonlinear Vibrations Using an Amplitude-dependent Autoregressive Time Series Model", *Biometrika*, 68, 189-196.

Hall, P. (1992), *The Bootstrap and Edgeworth Expansion,* Springer-Verlag.

Härdle, W. (1990), *Applied Nonparametric Regression*, New York, Cambridge University Press.

Härdle, W. and E. Mammen (1993), "Comparing Nonparametric versus Parametric Regression Fits," *Annals of Statistics*, 21, 1926-1947.

Hjellvik, V. and D. Tjøstheim (1995), "Nonparametric Tests of Linearity for Time Series", *Biometrika*, 82, 351-368.

Hjellvik, V. and D. Tjøstheim (1998), "Nonparametric Statistics for Testing of Linearity and Serial Independence", *Journal of Nonparametric Statistics*, 6, 223-251.

Hjellvik, V., Q. Yao, and D. Tjøstheim (1998), "Linearity Testing Using Local Polynomial Approximation", *Journal of Statistical Planning and Inference*, 68, 295-321.

Hong, Y. and H. White (1995), "Consistent Specification Testing via Nonparametric Series Regression", *Econometrica,* 63, 1133-1160.

Kreiss, J.-P., M.H. Neumann, and Q. Yao (1998), "Bootstrap Tests for Simple Structures in Nonparametric Time Series Regression", University of Kent at Canterbury.

Lavergne, P. and Q. Vuong (1996), "Nonparametric Selection of Regressors: the Nonnested Case", *Econometrica,* 64, 207-219.

Lee, B.-J. (1992) "A Heteroskedasticity Test Robust to Conditional Mean Misspecification", *Econometrica*, 60, 159-171.

Lee, T.-H., H. White and C.W.J. Granger (1993), "Testing for Neglected Nonlinearity in Time Series Models: A Comparison of Neural Network Methods and Alternative Tests", *Journal of Econometrics*, 56, 269-290.

Lewbel, A. (1993), "Consistent Tests with Nonparametric Components with an Application to Chinese Production Data", Brandeis University.

Lewbel, A. (1995), "Consistent Nonparametric Testing with an Application to Testing Slutsky Symmetry", *Journal of Econometrics,* 67, 379-401.

Li, Q. (1999), "Consistent Model Specification Tests for Time Series Econometric Models", *Journal of Econometrics,* 92(1), 101-147.

Li, Q., C.J. Huang, and T.-T. Fu (1997), "Semiparametric Smooth Coefficient Models", Texas A&M University.

Li, Q. and S. Wang (1998), "A Simple Consistent Bootstrap Test for a Parametric Regression Function", *Journal of Econometrics,* 87(1), 145-165.

Linton, O. and P. Gozalo (1997), "Consistent Testing of Additive Models".

Liu, R.Y. (1988), "Bootstrap Procedures under Some Non-iid Models", *Annals of Statistics*, 16, 1697-1708.

Nadaraya, É.A. (1964), "On Estimating Regression", *Theory of Probability and its Applications*, 9, 141-142.

Pagan, A.R. and A. Ullah (1999), *Nonparametric Econometrics*, Cambridge University Press.

Rahman, M. and A. Ullah (1999), "Improved Combined Parametric and Nonparametric Regressions: Estimation and Hypothesis Testing", UC Riverside.

Raj, B. and A. Ullah (1981), *Econometrics: A Varying Coefficients Approach*, Croom Helm, London.

Robinson, P.M. (1989), "Hypothesis Testing in Semiparametric and Nonparametric Models for Econometric Time Series", *Review of Economic Studies,* 56, 511-534.

Ruppert, D. and M.P. Wand (1994), "Multivariate Locally Weighted Least Squares Regression", *Annals of Statistics,* 22, 1346-1370.

Shao, J. and D. Tu (1995), *The Jackknife and Bootstrap*, Springer.

Stone, C.J. (1977), "Consistent Nonparametric Regression", *Annals of Statistics*, 5, 595-645.

Teräsvirta, T., C.-F. Lin and C.W.J. Granger (1993), "Power of the Neural Network Linearity Test", *Journal of Time Series Analysis,* 14, 209-220.

Tong, H. (1990), *Nonlinear Time Series: A Dynamic System Approach,* Clarendon Press, Oxford.

Ullah, A. (1985), "Specification Analysis of Econometric Models", *Journal of Quantitative Economics*, 1, 187-209.

Ullah, A. (1999), "Nonparametric Kernel Methods of Estimation and Hypothesis Testing", forthcoming, *Companion in Econometric Theory,* B. Baltagi (ed.), Basil Blackwell, U.K.

Ullah, A. and H. D. Vinod (1993), "General Nonparametric Regression Estimation and Testing in Econometrics", in G. S. Maddala, C. R. Rao and H. D. Vinod (eds.) *Handbook of Statistics*, Vol. 11, Amsterdam, Elseiver Publishing Co., 85-116.

Watson, G.S. (1964), "Smooth Regression Analysis", *Sankhya*, Series A, 26, 359-372.

Whang, Y.J. and D.W.K. Andrews (1993), "Testing of Specification for Parametric and Semiparametric Models", *Journal of Econometrics*, 57, 277-318.

Wooldridge, J.M. (1992), "A Test for Functional Form Against Nonparametric Alternatives", *Econometric Theory*, 8, 452-475.

Wu, C.F.J. (1986), "Jackknife, Bootstrap, and Other Resampling Methods in Regression Analysis", *Annals of Statistics*, 14, 1261-1350.

Yatchew, A.J. (1992), "Nonparametric Regression Tests Based on Least Squares", *Econometric Theory*, 8, 435-451.

Zheng, J.X. (1996), "A Consistent Test of Functional Form via Nonparametric Estimation Techniques", *Journal of Econometrics,* 75(2), 263-289.

TABLE 1. Size

Panel A 5% nominal level of significance

Block	DGP	n	T^B	T^W	$L^A_{0.5}$	$L^B_{0.5}$	$L^W_{0.5}$	$L^A_{1.0}$	$L^B_{1.0}$	$L^W_{1.0}$	$L^A_{2.0}$	$L^B_{2.0}$	$L^W_{2.0}$
1	AR	50	.044	.043	.013	.027	.030	.001	.018	.015	.000	.002	.002
		100	.034	.031	.015	.029	.027	.007	.020	.018	.001	.009	.012
		200	.028	.026	.023	.036	.036	.007	.026	.023	.001	.009	.010
2	M1	50	.046	.064	.021	.028	.027	.008	.022	.018	.000	.002	.002
		100	.015	.030	.028	.030	.035	.012	.028	.026	.000	.007	.006
		200	.024	.025	.030	.032	.045	.006	.018	.020	.000	.003	.007
2	M4	50	.047	.070	.022	.025	.026	.008	.019	.016	.000	.003	.004
		100	.023	.032	.025	.030	.029	.009	.018	.021	.000	.006	.006
		200	.014	.019	.034	.041	.037	.015	.031	.030	.001	.006	.004
4	Z1	50	.019	.131	.031	.060	.063	.010	.062	.057	.001	.060	.050
		100	.019	.099	.027	.044	.044	.013	.049	.047	.000	.053	.058
		200	.014	.117	.023	.043	.044	.011	.043	.045	.000	.045	.045

Panel B 10% nominal level of significance

Block	DGP	n	T^B	T^W	$L^A_{0.5}$	$L^B_{0.5}$	$L^W_{0.5}$	$L^A_{1.0}$	$L^B_{1.0}$	$L^W_{1.0}$	$L^A_{2.0}$	$L^B_{2.0}$	$L^W_{2.0}$
1	AR	50	.072	.090	.033	.063	.063	.005	.042	.044	.000	.010	.013
		100	.069	.072	.035	.075	.074	.010	.052	.048	.001	.023	.024
		200	.062	.057	.042	.085	.082	.013	.063	.059	.001	.033	.030
2	M1	50	.071	.122	.053	.075	.074	.014	.052	.044	.000	.013	.011
		100	.035	.073	.050	.072	.070	.025	.054	.054	.000	.017	.014
		200	.045	.063	.063	.087	.091	.021	.049	.058	.000	.021	.022
2	M4	50	.077	.112	.040	.061	.058	.016	.035	.035	.000	.013	.012
		100	.047	.069	.045	.068	.066	.020	.038	.039	.001	.018	.015
		200	.042	.040	.065	.085	.083	.030	.065	.063	.001	.018	.017
4	Z1	50	.044	.196	.051	.119	.118	.018	.110	.103	.001	.097	.099
		100	.043	.160	.043	.087	.090	.020	.113	.108	.000	.112	.115
		200	.045	.182	.045	.097	.093	.022	.087	.097	.001	.096	.106

Notes: Test statistics are denoted as T^j and L^j_c, with the superscripts $j = A, B, W$ refer to the methods of obtaining the null distributions of the test statistics; using the asymptotics (A), naive bootstrap (B), and wild bootstrap (W). The number of bootstrap resamples = 500 and number of monte carlo replications = 1000. The 95% asymptotic confidence interval of the estimated size is (.036, .064) at 5% nominal level of significance and (.081, .119) at 10% nominal level of significance.

TABLE 2. Size under conditional heteroskedasticity

Panel A 5% nominal level of significance

Block	DGP	n	T^B	T^W	$L^A_{0.5}$	$L^B_{0.5}$	$L^W_{0.5}$	$L^A_{1.0}$	$L^B_{1.0}$	$L^W_{1.0}$	$L^A_{2.0}$	$L^B_{2.0}$	$L^W_{2.0}$
1	AR′	50	.196	.164	.029	.048	.037	.010	.047	.024	.000	.022	.006
	$\alpha = .5$	100	.304	.179	.045	.061	.052	.025	.065	.042	.005	.058	.023
	$\beta = .0$	200	.450	.195	.033	.059	.042	.028	.074	.044	.010	.075	.034
1	AR′	50	.283	.234	.042	.071	.036	.025	.098	.037	.005	.076	.018
	$\alpha = .7$	100	.452	.267	.061	.108	.056	.051	.127	.053	.015	.120	.037
	$\beta = .0$	200	.662	.309	.074	.111	.073	.072	.137	.069	.034	.175	.059
1	AR′	50	.057	.068	.014	.026	.024	.002	.017	.014	.000	.006	.001
	$\alpha = .1$	100	.079	.060	.025	.038	.040	.013	.027	.027	.000	.011	.011
	$\beta = .89$	200	.165	.076	.043	.058	.055	.024	.056	.047	.004	.036	.021
1	AR′	50	.154	.147	.030	.051	.038	.012	.052	.024	.001	.026	.006
	$\alpha = .3$	100	.274	.141	.040	.071	.050	.020	.070	.037	.004	.049	.020
	$\beta = .69$	200	.568	.228	.059	.091	.053	.044	.118	.047	.016	.133	.037
1	AR′	50	.229	.190	.034	.057	.032	.023	.070	.036	.003	.045	.017
	$\alpha = .5$	100	.441	.242	.051	.092	.046	.039	.102	.043	.015	.112	.029
	$\beta = .49$	200	.711	.324	.093	.148	.070	.089	.189	.077	.046	.220	.060
4	Z1′	50	.264	.367	.048	.089	.053	.033	.117	.062	.006	.170	.065
		100	.382	.357	.040	.079	.045	.031	.127	.051	.011	.170	.055
		200	.506	.386	.046	.077	.055	.045	.125	.053	.019	.193	.051

Panel B 10% nominal level of significance

Block	DGP	n	T^B	T^W	$L^A_{0.5}$	$L^B_{0.5}$	$L^W_{0.5}$	$L^A_{1.0}$	$L^B_{1.0}$	$L^W_{1.0}$	$L^A_{2.0}$	$L^B_{2.0}$	$L^W_{2.0}$
1	AR′	50	.287	.241	.046	.104	.079	.019	.106	.068	.002	.056	.026
	$\alpha = .5$	100	.398	.258	.063	.119	.093	.043	.118	.075	.007	.104	.047
	$\beta = .0$	200	.548	.295	.067	.113	.086	.049	.124	.091	.020	.125	.073
1	AR′	50	.374	.308	.067	.131	.086	.047	.158	.082	.011	.123	.055
	$\alpha = .7$	100	.562	.363	.102	.183	.123	.072	.193	.102	.027	.182	.074
	$\beta = .0$	200	.747	.409	.108	.179	.122	.090	.237	.132	.054	.259	.120
1	AR′	50	.097	.117	.024	.066	.067	.005	.039	.041	.000	.013	.014
	$\alpha = .1$	100	.146	.108	.040	.068	.068	.018	.061	.052	.002	.029	.023
	$\beta = .89$	200	.247	.123	.062	.098	.092	.043	.091	.075	.009	.069	.051
1	AR′	50	.228	.204	.048	.105	.079	.022	.090	.058	.001	.047	.022
	$\alpha = .3$	100	.363	.213	.068	.123	.091	.034	.118	.070	.006	.093	.038
	$\beta = .69$	200	.653	.321	.091	.163	.101	.071	.185	.103	.032	.188	.081
1	AR′	50	.327	.271	.050	.126	.073	.035	.116	.072	.009	.081	.037
	$\alpha = .5$	100	.537	.330	.078	.160	.092	.061	.163	.072	.020	.162	.060
	$\beta = .49$	200	.776	.434	.137	.231	.142	.120	.265	.138	.066	.286	.112
4	Z1′	50	.372	.465	.085	.152	.112	.055	.209	.114	.011	.257	.138
		100	.483	.436	.077	.149	.098	.055	.207	.097	.017	.260	.109
		200	.614	.464	.081	.149	.106	.076	.196	.108	.027	.271	.120

Notes: AR′ is AR with GARCH(1,1) $h_t = (1 - \alpha - \beta) + \alpha\varepsilon^2_{t-1} + \beta h^2_{t-1}$. Z1′ is Z1 with $h_t \equiv E(\varepsilon^2_t|\xi_t) = (1 + x^2_{t1} + x^2_{t2})/3$.

TABLE 3 Power (5% level)

Block	DGP	n	T^B	T^W	$L^A_{0.5}$	$L^B_{0.5}$	$L^W_{0.5}$	$L^A_{1.0}$	$L^B_{1.0}$	$L^W_{1.0}$	$L^A_{2.0}$	$L^B_{2.0}$	$L^W_{2.0}$
1	BL	50	.380	.319	.043	.069	.040	.027	.084	.031	.003	.082	.015
		100	.616	.428	.062	.121	.044	.053	.162	.052	.026	.179	.042
1	TAR	50	.311	.329	.408	.482	.461	.228	.410	.398	.001	.069	.074
		100	.717	.645	.876	.902	.908	.791	.884	.881	.095	.527	.546
1	SGN	50	.392	.408	.536	.629	.621	.334	.529	.534	.005	.145	.143
		100	.838	.796	.962	.971	.972	.890	.955	.960	.183	.678	.710
1	NAR	50	.073	.098	.042	.062	.061	.020	.052	.047	.000	.021	.026
		100	.045	.048	.038	.057	.057	.019	.060	.057	.004	.036	.039
2	M2	50	.083	.121	.047	.060	.048	.015	.027	.025	.000	.010	.009
		100	.076	.120	.046	.060	.050	.030	.063	.050	.002	.027	.019
2	M3	50	.216	.238	.081	.101	.091	.089	.141	.125	.019	.157	.128
		100	.484	.448	.184	.211	.205	.270	.359	.346	.177	.479	.419
2	M5	50	.691	.644	.302	.344	.309	.276	.425	.362	.019	.255	.155
		100	.956	.893	.640	.680	.651	.749	.855	.798	.346	.802	.652
2	M6	50	.640	.618	.201	.242	.184	.115	.253	.153	.001	.124	.034
		100	.881	.789	.500	.540	.478	.466	.657	.474	.091	.449	.171
3	SQ	50	.303	.498	.172	.290	.250	.169	.418	.361	.050	.561	.476
		100	.703	.819	.425	.582	.535	.491	.725	.681	.355	.861	.815
3	EXP	50	.362	.499	.197	.294	.229	.197	.385	.294	.108	.481	.373
		100	.644	.758	.373	.477	.398	.407	.602	.507	.326	.728	.621
4	Z2	50	.977	.994	.773	.842	.697	.892	.972	.892	.863	.997	.974
		100	1.000	1.000	.979	.992	.968	.999	1.000	.995	1.000	1.000	1.000
4	Z3	50	.054	.255	.072	.134	.128	.056	.206	.195	.007	.283	.258
		100	.161	.410	.168	.255	.253	.193	.414	.411	.080	.541	.534
4	Z4	50	.999	1.000	.998	.999	.992	1.000	1.000	.996	1.000	1.000	1.000
		100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	EXPAR	50	.784	.755	.399	.358	.344	.305	.281	.259	.014	.068	.061
		100	.969	.958	.753	.748	.746	.798	.828	.810	.267	.515	.462
5	TAR	50	.209	.263	.119	.141	.131	.082	.143	.147	.002	.062	.059
		100	.389	.398	.203	.228	.224	.225	.317	.308	.050	.251	.239
6	LSTAR	50	.669	.679	.159	.156	.147	.044	.109	.094	.000	.013	.008
		100	.945	.920	.504	.512	.503	.340	.452	.425	.012	.158	.121
6	ESTAR	50	.316	.316	.098	.111	.106	.041	.061	.056	.001	.013	.008
		100	.615	.584	.259	.280	.289	.181	.255	.249	.008	.044	.039