CDE
March, 2000

Centre for Development Economics

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

Aman Ullah
Centre for Development Economics
&
Department of Economics
University of California
Riverside, CA 92521
ullah@mail.ucr.edu
Tae-Hwy Lee
Department of Economics
University of California
Riverside, CA 92521
taelee@mail.ucr.edu

Working Paper No. 77

Abstract

A unified framework for various nonparametric kernel regression estimators is 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.

JEL Classification: C12, C22

*We would like to thank Zongwu Cai, Jianqing Fan, and Qi Li for programs and discussion. 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 five decades 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 (1999), 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 $E(u_t|x_t) = 0$ and $Var(u_t|x_t) = \sigma^2$.

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. This $\hat{\delta}$ is obtained by minimizing

$$\sum u_t^2 = \sum (y_t - g(x_t, \delta))^2 \tag{2}$$

with respect to δ. For example, if $g(x_t, \delta) = X_t\delta$ is linear, we can obtain the LS estimator of δ as

$$\hat{\delta} = (X'X)^{-1}X'y, \tag{3}$$

and the predicted residuals $\hat{u}_t = y_t - \hat{m}(x_t)$, where X is an $n \times (k+1)$ matrix generated by $X_t = (1 \ \ x_t)$ and

$$\hat{m}(x_t) = X_t\hat{\delta} = X_t(X'X)^{-1}X'y. \tag{4}$$

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

$$\begin{aligned} y_t &= m(x_t) + u_t \\ &= m(x) + v_t \end{aligned} \tag{5}$$

where $v_t = (x_t - x)m^{(1)}(x) + \frac{1}{2}(x_t - x)^2 m^{(2)}(x) + \cdots + u_t$ 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

$$\sum_{t=1}^{n} v_t^2 K_{tx} = \sum_{t=1}^{n} (y_t - m(x))^2 K_{tx} \tag{6}$$

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

$$\tilde{m}(x) = \frac{\sum_{t=1}^{n} y_t K_{tx}}{\sum_{t=1}^{n} K_{tx}} = (\mathbf{i}' K(x) \mathbf{i})^{-1} \mathbf{i}' K(x) y \tag{7}$$

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

$$\tilde{m}(x) = \frac{\sum_{t'=1, t' \neq t}^{n} y_{t'} K_{t't}}{\sum_{t'=1, t' \neq t}^{n} K_{t't}}, \tag{8}$$

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 $Ev_t \to 0$ as $n \to \infty$. Thus the estimator $\tilde{m}(x)$ will be consistent under certain smoothing conditions on h, K, and $m(x)$. In small samples however $Ev_t \neq 0$ so $\tilde{m}(x)$ 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

$$\begin{aligned} y_t &= m(x_t) + u_t = m(x) + (x_t - x)m^{(1)}(x) + v_t \\ &= \alpha(x) + x_t \beta(x) + v_t \\ &= X_t \delta(x) + v_t \end{aligned} \tag{9}$$

where $\delta(x) = [\alpha(x)\ \beta(x)']'$ 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

$$\sum_{t=1}^{n} v_t^2 K_{tx} = \sum_{t=1}^{n} (y_t - X_t\delta(x))^2 K_{tx} \tag{10}$$

and it is given by

$$\tilde{\delta}(x) = (X'K(x)X)^{-1}X'K(x)y. \tag{11}$$

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

$$\tilde{m}(x) = [1\ x]\tilde{\delta}(x) = \tilde{\alpha}(x) + x\tilde{\beta}(x). \tag{12}$$

Obviously when $X = \mathrm{i}$, $\tilde{\delta}(x)$ reduces to the NW's LCLS estimator $\tilde{m}(x)$.

An estimator 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

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

with respect to $\delta(x)$. For $g(x_t, \delta(x)) = X_t\delta(x)$ we get the LLLS in (11). 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 $\sum[y_t - g(x_t, \delta(x))]^2$, that is the LLS becomes the global LS estimator given by (3).

The LLLS estimator in (11) can also be interpreted as the estimator of the functional coefficient (varying coefficient) linear regression model

$$\begin{aligned} y_t &= m(x_t) + u_t \\ &= X_t\delta(x) + u_t \end{aligned} \tag{14}$$

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 (11), 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

$$\begin{aligned} y_t &= m(x_t) + u_t = X_t\delta(x_t) + u_t \\ &\simeq X_t\delta(x) + X_tD(x)(x_t - x)' + u_t \end{aligned} \tag{15}$$

$$\begin{aligned} &= X_t\delta(x) + X_t \odot (x_t - x)vecD(x) + u_t \\ &= X_t^{*}\delta^*(x) + u_t \end{aligned}$$

where $X_t^* = [X_t \; X_t \odot (x_t - x)]$, $\delta^*(x) = [\delta'(x) \; (vecD(x))']'$, and $\odot$ represents the Hadamard product. The LL varying coefficient or LPLS estimator of $\delta^*(x)$ can then be obtained by minimizing

$$\sum_{t=1}^{n}[y_t - X_t^*\delta^*(x)]^2 K_{tx} \tag{16}$$

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

$$\tilde{\tilde{\delta}}^*(x) = (X^{*\prime}K(x)X^*)^{-1}X^{*\prime}K(x)y. \tag{17}$$

From this $\tilde{\tilde{\delta}}(x) = [I \; 0]\tilde{\tilde{\delta}}^*(x)$, and hence

$$\tilde{\tilde{m}}^*(x) = [1 \; x \; 0]\tilde{\tilde{\delta}}^*(x) = [1 \; x]\tilde{\tilde{\delta}}(x). \tag{18}$$

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

$$E(y_t|\xi_t) = m(\xi_t) = m(x_t, z_t) = X_t\delta(z_t), \tag{19}$$

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 random coefficient model (Raj and Ullah 1981, 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

$$\tilde{\delta}(z) = (X'K(z)X)^{-1}X'K(z)y \tag{20}$$

where $K(z)$ is a diagonal matrix of $K_{tz}, t = 1, \ldots, n$.

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

$$\begin{aligned} \sum_{t=1}^{n}[y_t - X_t\delta(z_t)]^2 K_{tz} &= \sum_{t=1}^{n}[y_t - X_t\delta(z) - (X_t \odot (z_t - z))vecD(z)]^2 K_{tz} \\ &= \sum_{t=1}^{n}[y_t - X_t^{**}\delta^*(z)]^2 K_{tz} \end{aligned} \tag{21}$$

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

$$\tilde{\tilde{\delta}}^{*}(z) = (X^{**\prime}K(z)X^{**})^{-1}X^{**\prime}K(z)y, \tag{22}$$

and $\tilde{\tilde{\delta}}(z) = [I \; 0]\tilde{\tilde{\delta}}^{*}(z)$. Hence

$$\tilde{\tilde{m}}^{*}(\xi) = [1 \; x]\tilde{\tilde{\delta}}(z). \tag{23}$$

When $z = x$, (20) and (22) reduce to the LLLS estimator $\tilde{\delta}(x)$ in (11) and the LL varying coefficient estimator $\tilde{\tilde{\delta}}^{*}(x)$ in (17), respectively. For the asymptotic properties of these varying coefficient estimators, see CFY (2000).

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

$$H_0 : P[E(y_t|\xi_t) = \xi_t\delta^{*}] = 1 \text{ for some } \delta^{*} \in \mathbb{R}^p. \tag{24}$$

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

$$H_1 : P[E(y_t|\xi_t) = \xi_t\delta] < 1 \text{ for all } \delta \in \mathbb{R}^p. \tag{25}$$

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$^{\text{P}}$), $\sum \hat{u}_t^2$, $\hat{u}_t = y_t - g(\xi_t, \hat{\delta})$ with the nonparametric RSS (RSS$^{\text{NP}}$), $\sum \tilde{u}_t^2$, where $\tilde{u}_t = y_t - \tilde{\tilde{m}}^{*}(\xi_t)$. The test statistic is

$$T = \frac{(\text{RSS}^{\text{P}} - \text{RSS}^{\text{NP}})}{\text{RSS}^{\text{NP}}} = \frac{\sum \hat{u}_t^2 - \sum \tilde{u}_t^2}{\sum \tilde{u}_t^2}, \tag{26}$$

or simply $T' = (\text{RSS}^{\text{P}} - \text{RSS}^{\text{NP}})$. 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 (1995) uses different normalizing factor and show the asymptotic normality of $nh^{p/2}T$.

Another way is to use the bootstrap method as suggested by CFY (2000). The bootstrap allows the implementation of (26) and it involves the following steps to evaluate p-values of the test:

1. Generate the bootstrap residuals $\{u_t^*\}$ from the centered NP residuals $(\tilde{u}_t - \bar{\tilde{u}})$ where $\bar{\tilde{u}} = n^{-1} \sum \tilde{u}_t$ and define $y_t^* \equiv \xi_t \hat{\delta} + u_t^*$.

2. Construct the bootstrap sample $\{y_t^*\ \xi_t\}_{t=1}^n$ and calculate the bootstrap test statistic T^* using, for the sake of simplicity, the same h used in estimation with the original sample.

3. Repeat the above two steps B times and use the empirical distribution of T^* as the null distribution of T. Reject the null hypothesis H_0 when T is greater than the upper α point of the conditional distribution of T^* given $\{y_t\ \xi_t\}_{t=1}^n$. The p-value of the test 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).

An intuitive and simple test of the parametric specification follows from the combined regression

$$y_t = g(\xi_t, \delta) + m_u(\xi_t) + \epsilon_t \tag{27}$$

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

$$E[u_t E(u_t|\xi_t) f(\xi_t)] = 0, \tag{28}$$

where $f(\xi_t)$ is the density of ξ. A sample estimator of the left hand side of (28) is

$$\begin{aligned} L' &= \frac{1}{n}\sum_{t=1}^{n} \hat{u}_t E(\hat{u}_t|\xi_t)\hat{f}(\xi_t) \\ &= \frac{1}{n(n-1)h^p}\sum_{t=1}^{n}\sum_{t'=1,t'\neq t}^{n} \hat{u}_t \hat{u}_{t'} K_{t't} \end{aligned} \tag{29}$$

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 (8) and $\hat{f}(\xi_t) = (nh^p)^{-1}\sum_{t'\neq t} K_{t't}$ is the kernel density estimator; $K_{t't} = K(\frac{\xi_{t'}-\xi_t}{h})$. The asymptotic test statistic is then given by

$$L = nh^{p/2}\frac{L'}{\sqrt{\hat{\omega}}} \sim N(0,1) \tag{30}$$

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.

3 Monte carlo

In this section we examine the finite sample properties of T and L especially with the empirical null distributions being generated by the bootstrap method. Asymptotic critical values are also used for L. 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 shown 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**)

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

where $\mathrm{sign}(x) = 1$ if $x > 0, 0$ if $x = 0$, and -1 if $x < 0$.

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 $x_t = (x_{t1}\ x_{t2})$ 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′**)

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

the

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, 1999)

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 (3) and $\hat{u}_t$ for the linear model, and (23) 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 T, 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

$$AMS(h) = \sum_{q=1}^{Q} AMS_q(h) \tag{31}$$

where

$$AMS_q(h) = \frac{1}{m} \sum_{t=n-qm+1}^{n-qm+m} [y_t - X_t \tilde{\tilde{\delta}}_q(z_t)]^2 \tag{32}$$

and $\tilde{\tilde{\delta}}_q(\cdot)$ 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 $z_t = x_{1t}$ for Block 4. For Block 5, $z_t = y_{t-1}$ for EXPAR and $z_t = y_{t-2}$ for TAR.

For L, 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 (29) and (30) is replaced with $\prod_{i=1}^{p} h_i$.

Test statistics are denoted as T^i and L_c^i, with the superscripts $i = A, B, W$ referring to the methods of obtaining the null distributions of the test statistics; asymptotics $(i = A)$, naive bootstrap $(i = B)$, and wild bootstrap $(i = W)$. Monte carlo experiments are conducted with 300 bootstrap resamples and 300 monte carlo replications.

Table 1 gives the estimated size of the tests for the data generating processes (DGP) which are linear in conditional mean. Table 1 reports the cases with the conditional homoskedastic errors. The 95% confidence interval of the estimated size is (0.025, 0.075) at 5% nominal level of significance, and (0.066, 0.134) at 10% nominal level of significance, since if the true size is s (e.g., $s = 0.05, 0.10$) the estimated size follows the asymptotic normal distribution with mean s and variance $s(1-s)/300$ with 300 monte carlo replications. Due to the long computing time for each simulation, only 300 replications are conducted and thus it must be noted that the confidence intervals are rather wide. The naive bootstrap CFY test T^B tends to under-reject the null, while T^W with wild bootstrap method tends to over-reject the null. The size is often worse with $n = 50$, which may be due to an estimation of the bootstrap DGP (23) to generate the bootstrap residuals $\{u_t^*\}$ in the very small samples. While the two bootstrap procedures work differently for the CFY test (T^B and T^W), they are very similar for the LWZ test (L_c^B and L_c^W). Both bootstrap tests L_c^B and L_c^W are generally better than the asymptotic test L_c^A. Both bootstrap procedures work very well especially with $c = 0.5$. $L_{0.5}$ is better than $L_{1.0}$ which is better than $L_{2.0}$. The size of L is quite sensitive to the choice of c and the bandwidth h.

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. A further refinement, beyond $n^{-1/2}$, of the order $n^{-1/2}$ can be obtained when an asymptotically pivotal statistic (whose limiting distribution is independent of unknown nuisance parameters) is used for testing. 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 by a full order of n^{-1}. See Hall (1992) for further discussion based on Edgeworth expansions on the extent of the refinements in other 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 (24) which will then imply equation (28) 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. However, the results in Table 2 show that the size of 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 condition. T is constructed to compare the two residual sums of squares RSS^{P} and RSS^{NP}. As the alternative model to compute RSS^{NP} is estimated by the functional coefficient (FC) model (23), 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. However, 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. 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 good 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 for neglected nonlinearity in regression models. Both naive bootstrap and wild bootstrap are used to generate the critical values together with the asymptotic distributions. T^B with the naive bootstrap tends to under-reject the null, while T^W with wild bootstrap method tends to over-reject the null. The bootstrap LWZ tests L^B and L^W are better than the asymptotic test L^A. When the errors are conditionally heteroskedastic the wild bootstrap for the LWZ test corrects the size distortion. 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 forecast improvement via a nonparametric model over a linear model. Hence, L 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.

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. Douglas, and Dingding 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, Tim (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics,* 31, 307-327.

Cai, Zongwu, Jianqing Fan, and Qiwei 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.

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

Davidson, Russell and James 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.

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 (1995), "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", forthcoming, *Journal of Nonparametric Statistics.*

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, Clive W.J. and Timo Teräsvirta (1993), *Modelling Nonlinear Economic Relationships,* Oxford University Press, New York.

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

Hall, Peter (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 (1999), "Linearity Testing Using Local Polynomial Approximation", *Journal of Statistical Planning and Inference*, forthcoming.

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

Kreiss, Jens-Peter, Michael H. Neumann, and Qiwei 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, Byung-Joo (1992) "A Heteroskedasticity Test Robust to Conditional Mean Misspecification", *Econometrica*, 60, 159-171.

Lee, Tae-Hwy, Halbert White and Clive 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, Qi (1999), "Consistent Model Specification Tests for Time Series Econometric Models", *Journal of Econometrics*, 92(1), 101-147.

Li, Qi and S. Wang (1998), "A Simple Consistent Bootstrap Test for a Parametric Regression Function", *Journal of Econometrics*, 87, 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, Timo, Chien-Fu Lin and Clive 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, Aman (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, John Xu (1996), "A Consistent Test of Functional Form via Nonparametric Estimation Techniques", *Journal of Econometrics,* 75, 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	.037	.127	.023	.033	.043	.007	.030	.033	.000	.003	.007
		100	.023	.087	.010	.017	.017	.003	.020	.023	.000	.007	.003
		200	.040	.097	.017	.033	.040	.007	.020	.023	.000	.003	.000
		300	.037	.073	.023	.033	.027	.010	.027	.020	.000	.010	.010
2	M1	50	.027	.170	.030	.033	.037	.013	.020	.027	.000	.000	.000
		100	.007	.113	.030	.033	.040	.010	.023	.020	.000	.000	.003
		200	.023	.100	.053	.053	.053	.013	.033	.030	.000	.000	.000
		300	.027	.093	.033	.037	.037	.020	.020	.023	.000	.007	.010
2	M4	50	.003	.113	.023	.027	.023	.010	.013	.017	.000	.000	.003
		100	.010	.100	.020	.023	.023	.010	.017	.013	.000	.007	.007
		200	.020	.097	.023	.030	.027	.017	.027	.017	.000	.017	.017
		300	.017	.110	.033	.040	.033	.007	.013	.013	.000	.003	.003
4	Z1	50	.023	.117	.033	.073	.073	.003	.040	.047	.000	.030	.033
		100	.007	.077	.020	.050	.050	.017	.057	.070	.000	.063	.060
		200	.023	.097	.027	.040	.050	.010	.047	.037	.003	.063	.050
		300	.017	.080	.017	.037	.030	.003	.037	.030	.000	.040	.047

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	.080	.193	.040	.083	.073	.020	.057	.053	.000	.027	.023
		100	.053	.147	.023	.063	.060	.020	.057	.047	.000	.027	.027
		200	.060	.143	.053	.090	.083	.017	.060	.050	.000	.020	.023
		300	.067	.143	.033	.070	.077	.013	.047	.050	.000	.040	.040
2	M1	50	.050	.220	.067	.087	.093	.020	.043	.043	.000	.013	.007
		100	.037	.177	.053	.073	.073	.023	.057	.060	.000	.017	.020
		200	.037	.153	.077	.090	.090	.033	.063	.053	.000	.003	.017
		300	.047	.157	.050	.080	.073	.027	.057	.050	.000	.023	.023
2	M4	50	.013	.163	.040	.060	.053	.013	.043	.040	.000	.010	.010
		100	.037	.150	.030	.043	.057	.013	.033	.037	.003	.013	.010
		200	.047	.157	.033	.057	.050	.023	.040	.043	.020	.020	.020
		300	.033	.187	.057	.073	.067	.013	.043	.033	.000	.013	.020
4	Z1	50	.053	.190	.080	.137	.130	.010	.130	.123	.000	.077	.107
		100	.030	.160	.047	.130	.130	.027	.110	.103	.000	.113	.110
		200	.043	.170	.050	.100	.103	.017	.080	.087	.003	.120	.107
		300	.040	.147	.037	.083	.073	.010	.090	.090	.000	.093	.110

Notes: Test statistics are denoted as T^i and L^i_c, with the superscripts $i = A, B, W$ 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 = 300 and number of monte carlo replications = 300.

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	.183	.230	.040	.067	.043	.013	.063	.040	.003	.043	.017
	$\alpha = .5$	100	.260	.290	.030	.063	.040	.013	.067	.043	.000	.067	.027
	$\beta = .0$	200	.417	.373	.040	.067	.043	.020	.070	.033	.003	.057	.020
1	AR′	50	.240	.337	.020	.047	.023	.020	.053	.023	.000	.040	.013
	$\alpha = .7$	100	.403	.430	.047	.067	.047	.040	.087	.047	.027	.090	.037
	$\beta = .0$	200	.610	.520	.063	.127	.060	.060	.163	.053	.037	.177	.043
1	AR′	50	.043	.150	.023	.037	.030	.010	.033	.027	.000	.003	.010
	$\alpha = .1$	100	.063	.123	.010	.027	.020	.003	.020	.017	.000	.013	.013
	$\beta = .89$	200	.073	.110	.020	.030	.027	.007	.020	.017	.000	.023	.010
1	AR′	50	.137	.217	.007	.027	.023	.003	.027	.013	.000	.007	.003
	$\alpha = .3$	100	.233	.273	.043	.073	.043	.037	.083	.043	.003	.083	.030
	$\beta = .69$	200	.427	.303	.080	.123	.077	.063	.127	.060	.023	.140	.053
1	AR′	50	.167	.273	.043	.060	.040	.023	.060	.030	.000	.040	.017
	$\alpha = .5$	100	.380	.323	.053	.103	.043	.030	.090	.033	.007	.087	.017
	$\beta = .49$	200	.683	.527	.103	.170	.073	.093	.223	.083	.040	.253	.063
4	Z1′	50	.267	.387	.030	.063	.043	.020	.103	.043	.010	.117	.050
		100	.410	.360	.047	.080	.060	.027	.137	.060	.013	.170	.070
		200	.507	.350	.047	.083	.047	.020	.090	.040	.013	.163	.050

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	.233	.310	.063	.127	.087	.037	.093	.073	.003	.070	.037
	$\alpha = .5$	100	.377	.370	.050	.097	.070	.037	.117	.067	.010	.110	.050
	$\beta = .0$	200	.473	.453	.067	.123	.110	.037	.127	.097	.020	.110	.060
1	AR′	50	.330	.407	.043	.127	.087	.027	.090	.053	.003	.073	.033
	$\alpha = .7$	100	.547	.493	.063	.120	.093	.060	.153	.083	.030	.153	.063
	$\beta = .0$	200	.713	.590	.120	.190	.120	.097	.203	.113	.053	.237	.097
1	AR′	50	.073	.220	.037	.070	.060	.017	.050	.043	.000	.027	.020
	$\alpha = .1$	100	.100	.180	.023	.047	.053	.007	.053	.053	.003	.040	.033
	$\beta = .89$	200	.120	.160	.030	.073	.063	.010	.053	.050	.000	.030	.020
1	AR′	50	.213	.273	.030	.097	.063	.003	.083	.043	.000	.037	.007
	$\alpha = .3$	100	.340	.340	.077	.147	.083	.050	.157	.087	.013	.133	.067
	$\beta = .69$	200	.513	.397	.110	.200	.140	.087	.210	.113	.047	.200	.087
1	AR′	50	.273	.317	.057	.107	.070	.023	.103	.057	.003	.057	.027
	$\alpha = .5$	100	.447	.403	.083	.163	.100	.050	.170	.067	.010	.150	.053
	$\beta = .49$	200	.757	.617	.150	.280	.140	.147	.303	.153	.070	.337	.130
4	Z1′	50	.373	.450	.053	.170	.097	.030	.220	.083	.010	.243	.097
		100	.513	.480	.087	.153	.103	.063	.237	.120	.013	.263	.133
		200	.593	.460	.087	.150	.117	.047	.177	.090	.027	.247	.103

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 $h_t \equiv E(\varepsilon_t^2|\xi_t) = (1 + x^2_{t1} + x^2_{t2})/3$.

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	.357	.373	.060	.103	.053	.050	.120	.033	.010	.097	.017
		100	.600	.520	.073	.127	.037	.077	.170	.050	.047	.187	.040
		200	.773	.613	.130	.187	.107	.133	.230	.123	.110	.243	.080
1	TAR	50	.300	.567	.433	.503	.487	.220	.417	.400	.000	.090	.097
		100	.670	.827	.853	.880	.883	.743	.867	.880	.073	.487	.497
		200	.987	.993	.997	.997	.997	.997	.997	.997	.797	.977	.973
1	SGN	50	.397	.607	.533	.600	.603	.290	.513	.497	.007	.150	.140
		100	.887	.933	.970	.977	.977	.877	.960	.960	.150	.687	.743
		200	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.957	.997	.997
1	NAR	50	.023	.167	.017	.033	.023	.003	.033	.023	.000	.013	.003
		100	.057	.147	.043	.083	.067	.033	.097	.090	.003	.043	.053
		200	.080	.213	.087	.117	.123	.053	.130	.117	.010	.087	.090
2	M2	50	.070	.230	.033	.050	.040	.017	.027	.020	.000	.017	.020
		100	.073	.250	.037	.040	.043	.033	.060	.050	.003	.037	.017
		200	.077	.257	.090	.110	.103	.047	.097	.083	.010	.050	.030
2	M3	50	.227	.433	.077	.097	.100	.097	.177	.157	.027	.190	.143
		100	.497	.663	.210	.220	.227	.313	.390	.350	.190	.480	.430
		200	.857	.910	.433	.450	.447	.677	.733	.713	.650	.853	.833
2	M5	50	.627	.790	.273	.323	.263	.247	.427	.303	.017	.220	.097
		100	.973	.950	.607	.643	.610	.740	.853	.780	.353	.817	.590
		200	1.000	.997	.927	.940	.930	.993	.997	.993	.970	.993	.980
2	M6	50	.560	.670	.237	.283	.230	.137	.290	.147	.010	.147	.027
		100	.877	.820	.513	.580	.463	.500	.673	.467	.083	.480	.167
		200	.993	.933	.857	.883	.810	.913	.973	.913	.507	.907	.607
3	SQ	50	.350	.540	.203	.307	.280	.200	.407	.347	.063	.533	.483
		100	.727	.817	.430	.563	.530	.453	.700	.650	.363	.847	.810
		200	.980	.990	.833	.887	.873	.873	.970	.943	.840	.997	.990
3	EXP	50	.320	.463	.163	.247	.200	.173	.363	.273	.087	.467	.363
		100	.687	.767	.377	.497	.410	.423	.617	.527	.350	.713	.607
		200	.927	.957	.657	.733	.677	.723	.860	.807	.703	.933	.907

(Table 3 continued)

Block	DGP	n	T^B	T^W	$L^A_{0.5}$	$L^B_{0.5}$	$L^W_{0.5}$	$L^A_{1.0}$	$L^B_{1.0}$	$L^W_{1.0}$	$L^A_{2.0}$	$L^B_{2.0}$	$L^W_{2.0}$
4	Z2	50	.963	.990	.727	.813	.703	.860	.957	.863	.827	.983	.940
		100	1.000	1.000	.990	.997	.980	.997	1.000	.997	.997	1.000	.997
		200	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	Z3	50	.097	.257	.080	.150	.133	.077	.223	.213	.023	.263	.260
		100	.163	.383	.160	.223	.213	.207	.383	.373	.067	.533	.493
		200	.423	.633	.327	.400	.433	.467	.650	.673	.353	.793	.793
4	Z4	50	1.000	1.000	1.000	1.000	.990	1.000	1.000	1.000	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
		200	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	.613	.780	.407	.427	.407	.270	.370	.337	.010	.110	.097
		100	.953	.983	.763	.773	.767	.783	.837	.830	.280	.603	.520
		200	1.000	1.000	.993	.993	.987	1.000	1.000	1.000	.950	.990	.993
5	TAR	50	.127	.373	.103	.113	.127	.087	.123	.137	.000	.057	.063
		100	.433	.713	.247	.283	.293	.263	.353	.350	.067	.273	.280
		200	.850	.927	.543	.550	.553	.683	.753	.727	.467	.777	.763
6	LSTAR	50	.403	.513	.170	.257	.227	.040	.157	.127	.003	.030	.013
		100	.880	.933	.493	.560	.543	.327	.563	.500	.010	.210	.137
		200	1.000	1.000	.910	.940	.913	.900	.967	.953	.360	.833	.787
6	ESTAR	50	.143	.237	.100	.117	.110	.040	.093	.073	.003	.013	.013
		100	.543	.710	.280	.317	.313	.203	.277	.273	.003	.047	.047
		200	.943	.940	.620	.643	.647	.647	.683	.700	.103	.297	.297

CENTRE FOR DEVELOPMENT ECONOMICS
WORKING PAPER SERIES

No.	Author(s)	Title
1	Kaushik Basu Arghya Ghosh Tridip Ray	The Babu and The Boxwallah : Managerial Incentives and Government Intervention (January 1994). Review of Development Economics, 1997
2	M.N. Murty Ranjan Ray	Optimal Taxation and Resource Transfers in a Federal Nation (February 1994)
3	V. Bhaskar Mushtaq Khan	Privatization and Employment : A Study of The Jute Industry in Bangladesh (March 1994). American Economic Review, March 1995, pp. 267-273
4	V. Bhaskar	Distributive Justice and The Control of Global Warming (March 1994) The North, the South and the Environment: V. Bhaskar and Andrew Glyn (Ed.) Earthscan Publication London, February 1995
5	Bishnupriya Gupta	The Great Depression and Brazil's Capital Goods Sector: A Re-examination (April 1994). Revista Brasileria de Economia 1997
6	Kaushik Basu	Where There Is No Economist: Some Institutional and Legal Prerequisites of Economic Reform in India (May 1994)
7	Partha Sen	An Example of Welfare Reducing Tariff Under Monopolistic Competition (May 1994), Reveiw of International Economics, (forthcoming)
8	Partha Sen	Environmental Policies and North-South Trade : A Selected Survey of the Issues (May 1994)
9	Partha Sen Arghya Ghosh Abheek Barman	The Possibility of Welfare Gains with Capital Inflows in A Small Tariff-Ridden Economy (June 1994)
10	V. Bhaskar	Sustaining Inter-Generational Altruism when Social Memory is Bounded (June 1994)
11	V. Bhaskar	Repeated Games with Almost Perfect Monitoring by Privately Observed Signals (June 1994)

No.	Author(s)	Title
12	S. Nandeibam	Coalitional Power Structure in Stochastic Social Choice Functions with An Unrestricted Preference Domain (June 1994). Journal of Economic Theory (Vol. 68 No. 1, January 1996, pp. 212-233
13	Kaushik Basu	The Axiomatic Structure of Knowledge And Perception (July 1994)
14	Kaushik Basu	Bargaining with Set-Valued Disagreement (July 1994). Social Choice and Welfare, 1996, (Vol. 13, pp. 61-74)
15	S. Nandeibam	A Note on Randomized Social Choice and Random Dictatorships (July 1994). Journal of Economic Theory, Vol. 66, No. 2, August 1995, pp. 581-589
16	Mrinal Datta Chaudhuri	Labour Markets As Social Institutions in India (July 1994)
17	S. Nandeibam	Moral Hazard in a Principal-Agent(s) Team (July 1994) Economic Design Vol. 1, 1995, pp. 227-250
18	D. Jayaraj S. Subramanian	Caste Discrimination in the Distribution of Consumption Expenditure in India: Theory and Evidence (August 1994)
19	K. Ghosh Dastidar	Debt Financing with Limited Liability and Quantity Competition (August 1994)
20	Kaushik Basu	Industrial Organization Theory and Developing Economies (August 1994). Indian Industry: Policies and Performance, D. Mookherjee (ed.), Oxford University Press, 1995
21	Partha Sen	Immiserizing Growth in a Model of Trade with Monopolisitic Competition (August 1994). The Review of International Economics, (forthcoming)
22	K. Ghosh Dastidar	Comparing Cournot and Bertrand in a Homogeneous Product Market (September 1994)
23	K. Sundaram S.D. Tendulkar	On Measuring Shelter Deprivation in India (September 1994)
24	Sunil Kanwar	Are Production Risk and Labour Market Risk Covariant? (October 1994)

No.	Author(s)	Title
25	Partha Sen	Welfare-Improving Debt Policy Under Monopolistic Competition (November 1994)
26	Ranjan Ray	The Reform and Design of Commodity Taxes in the presence of Tax Evasion with Illustrative Evidence from India (December 1994)
27	Wietze Lise	Preservation of the Commons by Pooling Resources, Modelled as a Repeated Game (January 1995)
28	Jean Drèze Anne-C. Guio Mamta Murthi	Demographic Outcomes, Economic Development and Women's Agency (May 1995). Population and Development Review, December, 1995
29	Jean Drèze Jackie Loh	Literacy in India and China (May 1995). Economic and Political Weekly, 1995
30	Partha Sen	Fiscal Policy in a Dynamic Open-Economy New-Keynesian Model (June 1995)
31	S.J. Turnovsky Partha Sen	Investment in a Two-Sector Dependent Economy (June 1995). The Journal of Japanese and International Economics, Vol. 9, No. 1, March 1995
32	K. Krishnamurty V. Pandit	India's Trade Flows: Alternative Policy Scenarios: 1995-2000 (June 1995). Indian Economic Review, Vol. 31, No. 1, 1996
33	Jean Drèze P.V. Srinivasan	Widowhood and Poverty in Rural India: Some Inferences from Household Survey Data (July 1995). Journal of Development Economics, 1997
34	Ajit Mishra	Hierarchies, Incentives and Collusion in a Model of Enforcement (January 1996)
35	Sunil Kanwar	Does the Dog wag the Tail or the Tail the Dog? Cointegration of Indian Agriculture with Non-Agriculture (February 1996)
36	Jean Drèze P.V. Srinivasan	Poverty in India: Regional Estimates, 1987-8 (February 1996)
37	Sunil Kanwar	The Demand for Labour in Risky Agriculture (April 1996)
38	Partha Sen	Dynamic Efficiency in a Two-Sector Overlapping Generations Model (May 1996)

No.	Author(s)	Title
39	Partha Sen	Asset Bubbles in a Monopolistic Competitive Macro Model (June 1996)
40	Pami Dua Stephen M. Miller David J. Smyth	Using Leading Indicators to Forecast US Home Sales in a Bayesian VAR Framework (October 1996)
41	Pami Dua David J. Smyth	The Determinants of Consumers' Perceptions of Buying Conditions for Houses (November 1996)
42	Aditya Bhattacharjea	Optimal Taxation of a Foreign Monopolist with Unknown Costs (January 1997)
43	M. Datta-Chaudhuri	Legacies of the Independence Movement to the Political Economy of Independent India (April 1997)
44	Suresh D. Tendulkar T. A. Bhavani	Policy on Modern Small Scale Industries: A Case of Government Failure (May 1997)
45	Partha Sen	Terms of Trade and Welfare for a Developing Economy with an Imperfectly Competitive Sector (May 1997)
46	Partha Sen	Tariffs and Welfare in an Imperfectly Competitive Overlapping Generations Model (June 1997)
47	Pami Dua Roy Batchelor	Consumer Confidence and the Probability of Recession: A Markov Switching Model (July 1997)
48	V. Pandit B. Mukherji	Prices, Profits and Resource Mobilisation in a Capacity Constrained Mixed Economy (August 1997)
49	Ashwini Deshpande	Loan Pushing and Triadic Relations (September 1997)
50	Rinki Sarkar	Depicting the Cost Structure of an Urban Bus Transit Firm (September 1997)
51	Sudhir A. Shah	Existence and Optimality of Mediation Schemes for Games with Communication (November 1997)
52	V. Pandit	A Note on Data Relating to Prices in India (November 1997)
53	Rinki Sarkar	Cost Function Analysis for Transportation Modes: A Survey of Selective Literature (December 1997)

No.	Author(s)	Title
54	Rinki Sarkar	Economic Characteristics of the Urban Bus Transit Industry: A Comparative Analysis of Three Regulated Metropolitan Bus Corporations in India (February 1998)
55	Aditya Bhattacharjea	Was Alexander Hamilton Right? Limit-pricing Foreign Monopoly and Infant-industry Protection (February 1998)
56	Bishwanath Goldar Badal Mukherji	Pollution Abatement Cost Function: Methodological and Estimation Issues (March 1998)
57	Smita Misra	Economies of Scale in Water Pollution Abatement: A Case of Small-Scale Factories in an Industrial Estate in India (April 1998)
58	T.A. Bhavani Suresh D. Tendulkar	Determinants of Firm-level Export Performance: A Case Study of Indian Textile Garments and Apparel Industry (May 1998)
59	Partha Sen	Non-Uniqueness In The First Generation Balance of Payments Crisis Models (December 1998)
60	Ranjan Ray J.V. Meenakshi	State-Level Food Demand in India: Some Evidence on Rank-Three Demand Systems (December 1998)
61	Brinda Viswanathan	Structural Breaks in Consumption Patterns: India, 1952-1991 (December 1998)
62	Pami Dua Aneesa I. Rashid	Foreign Direct Investment and Economic Activity in India (March 1999)
63	Pami Dua Tapas Mishra	Presence of Persistence in Industrial Production : The Case of India (April 1999)
64	Sumit Joshi Sanjeev Goyal	Collaboration and Competition in Networks (May 1999)
65	Abhijit Banerji	Sequencing Strategically: Wage Negotiations Under Oligopoly (May 1999)
66	Jean Dreze Reetika Khera	Crime, Gender and Society in India : Some Clues from Homicide Data (June 1999)
67	Sumit Joshi	The Stochastic Turnpike Property without Uniformity in Convex Aggregate Growth Models (June 1999)

No.	Author(s)	Title
68	J.V. Meenakshi Ranjan Ray	Impact of Household Size, Family Composition and Socio Economic Characteristics on Poverty in Rural India (July 1999)
69	Jean Dreze Geeta Gandhi-Kingdon	School Participation in Rural India (September 1999)
70	Mausumi Das	Optimal Growth with Variable Rate of Time Preference (December 1999)
71	Chander Kant	Local Ownership Requirements and Total Tax Collections (January 2000)
72	K.N. Murty	Effects of Changes in Household Size, Consumer Taste & Preferences on Demand Pattern in India (January 2000)
73	Pami Dua Anirvan Banerji	An Index of Coincident Economic Indicators for the Indian Economy (January 2000)
74	V. Pandit	Macroeconometric Policy Modeling for India: A Review of Some Analytical Issues (February 2000)
75	A. Banerji	Selling Jointly Owned Assets via Bilateral Bargaining Procedures (March 2000)
76	Jean Dreze Mamta Murthi	Fertility, Education and Development: Further Evidence from India (March 2000)
77	Aman Ullah Tae-Hwy Lee	Nonparametric Bootstrap Tests for Neglected Nonlinearity in Time Series Regression Models (March 2000)