LM-Tests for Linearity Against Smooth Transition Alternatives: A Bootstrap Simulation Study*

Jonathan B. Hill†
Dept. of Economics
Florida International University

July 5, 2004

Abstract

The universal method for testing linearity against smooth transition autoregressive (STAR) alternatives is the linearization of the STAR model around the null nuisance parameter value, and performing F-tests on polynomial regressions in the spirit of the RESET test. Polynomial regressors, however, are poor proxies for the nonlinearity associated with STAR processes, and are not consistent (asymptotic power of one) against STAR alternatives, let alone general deviations from the null. Moreover, the most popularly used STAR forms of nonlinearity, exponential and logistic, are known to be exploitable for consistent conditional moment tests of functional form, cf. Bierens and Ploberger (1997). In this paper, pushing asymptotic theory aside, we compare the small sample performance of the standard polynomial test with an essentially ignored consistent conditional moment test of linear autoregression against smooth transition alternatives. In particular, we compute an LM sup-statistic and characterize the asymptotic p-value by Hansen's (1996) bootstrap method. In our simulations, we randomly select all STAR parameters in order not to bias experimental results based on the use of "safe", "interior" parameter values that exaggerate the smooth transition nonlinearity. Contrary to past studies, we find that the traditional polynomial regression method performs only moderately well, and that the LM sup-test out-performs the traditional test method, in particular for small samples and for LSTAR processes.

1. Introduction Smooth Threshold Autoregressive (STAR) models have gained significant popularity in the economics and finance literatures as a means

Key words: smooth transition autoregression; consistent conditional moment; bootstrap; simulation

†Dept. of Economics, Florida International University, Miami, Fl; Fax: 305-348-1524; jonathan.hill@fiu.edu; http://www.fiu.edu/~hilljona

to transcend well known estimation and forecasting limitations of both linear and binary switching (e.g. SETAR and Markov Switching) models. Original theoretical contributions belong to Tong (1983) and Chan and Tong (1986a,b), while Luukkonen *et al* (1988) and Teräsvirta (1994) develops a composite theory of estimation and testing for STAR processes with exponential and logistic transition functions. Models of smooth regime switching have been widely applied to exchange rates, prices, and stock returns, and the general theory has been extended to the GARCH class of conditional volatility, and models of flexible parametric form. See Teräsvirta (1994) and van Dijk *et al* (2000) for extensive bibliographies.

Consider a time series process $\{y_t\}$, regressors $x_{it} = (1, y_{t-1}, ..., y_{t-p_i})'$, $i = 1, 2$, and a stochastic shock u_t. The class of two-regime STAR processes is represented as

$$y_t = \phi_1' x_{1t} + \phi_2' x_{2t} F(y_{t-d}, \gamma, c) + u_t, \tag{1}$$

for some transition function $F_t(d, \gamma, c) = F(y_{t-d}, \gamma, c) : \mathbb{R}^3 \to [0, 1]$, transition scale $\gamma > 0$, threshold variable y_{t-d}, threshold c, and delay parameter d. The transition function is assumed to be twice continuously differentiable in γ and c. Luukkonen *et al* (1988), Teräsvirta (1994), and evidently the vast majority of applied research, consider logistic and exponential transition functions, respectively

$$F_t(d, \gamma, c) = \frac{1}{1 + e^{-\gamma(y_{t-d} - c)}}, \quad F_t(d, \gamma, c) = 1 - e^{-\gamma(y_{t-d} - c)^2}. \tag{2}$$

If the scale parameter γ and/or vector ϕ_2 are zero, then the process collapses to a linear autoregression.

Tests for linearity against STAR alternatives, however, have received little attention, and to date there do not exist treatments of consistent (asymptotic power of one) test methods. Under the traditional null hypothesis of linearity, $\gamma = 0$, the coefficients ϕ_2 are unidentified, and therefore standard Lagrange Multiplier statistics cannot be directly computed. Luukkonen *et al* (1988), Saikkonen and Luukkonen (1988), Teräsvirta (1994), Hagerud (1997), Gonzalez-Rivera (1998), Escribano and Jorda (2000), Madieros and Veiga (2000) and others proscribe a truncated Taylor expansion approximation of the nonlinear transition function $F_t(d, \gamma, c)$ around $\gamma = 0$ as a means to transcend the nuisance parameter and non-standard distribution dilemma. The technique leads to a simple polynomial auxiliary regression in the spirit of the RESET tests by Ramsey and Schmidt (1976) and Keenan (1985), and standard F-tests of parametric zero-restrictions. Tests on subsets of coefficients can be used to infer whether the process is exponential or logistic STAR. The simplicity of the auxiliary regression makes this method employable in any standard econometrics software and therefore has appeal for quick applications.

Several fundamental problems associated with polynomial regressions exist, however. First, by construction the resulting F-tests do not necessarily lead to a STAR model when the null of linearity is rejected, *unless it is assumed a priori* that the true data generating structure is STAR: see Teräsvirta (1994). The polynomial test amounts to a test of linearity on an assumed STAR process, and is not, therefore, a true test of smooth transition nonlinearity. To date, there does not exist a test which can reveal whether STAR nonlinearity provides a better approximation to the true data generating structure than the null specification. Pending evidence in favor of a smooth transition structure improving model fit, the polynomial regression would only then be appropriate for ascertaining which STAR model, exponential or logistic, best describes the data.

The polynomial regression technique only provides maximal power against local polynomial alternatives. This issue is particularly relevant if we admit *any* functional alternative to explain the data provided linearity is found inadequate, and are willing to use smooth transition nonlinearity to improve model performance[1]. Indeed, polynomial nonlinearity is known *not* to be "generically comprehensive" in the sense that if linearity is incorrect, additive polynomial terms may not improve the model fit: see Stinchcombe and White (1998). This shortcoming of classic weight-based moment condition specification tests is well known in the inference theory and artificial neural network literatures: see, e.g., Holley (1982), Davies (1987), Bierens (1990), and Kuan and White (1994).

Interestingly, the very nonlinear forms popularly espoused in the smooth transition literature, exponential and logistic, *are* generically comprehensive[2]. Orthogonality tests (e.g. a score test) which incorporate such functional weights have been shown to obtain asymptotic power of one against arbitrary deviations from the null functional specification: see Bierens (1982, 1990), Bierens and Ploberger (1997), and Stinchcombe and White (1998). While the laudable property of such structure absorption has been long recognized in the applied neural network literature, it has evidently been ignored entirely in the STAR literature. Indeed, under general conditions it is straightforward to show if $F_t(\cdot)$ is generically comprehensive, then so is $x_t F_t(\cdot)$. Thus, whereas a Bierens-type test employs the scalar weight $F_t(\cdot)$ and leads to a neural-network type model when the null is rejected, use of the weight $x_t F_t(\cdot)$ in is identically generically comprehensive and leads to a smooth transition type model when the null is rejected.

Second, for STAR tests based on one threshold variable y_{t-d}, the "delay" parameter d still exists in the polynomial regression. The delay parameter

[1] This is precisely the spirit in which the consistent parametric tests developed by Bierens (1990), Lee *et al* (1993) and Bierens and Ploberger (1997) are employed in neural network modeling. A null specification is tested without a prior alternative in mind, while rejection leads to a model with an additional neural network term that is guaranteed asymptotically to improve the model fit with probability one. See also Kuan and White (1994) and Stinchcombe and White (1998). The Hausman, RESET and McLeod-Li tests are well known examples of tests of model specification which are not consistent against general deviations, while consistent nonparametric tests typically do not provide a parametric alternative when the null model is found to be mis-specified: see, e.g., Yatchew (1992), Wooldridge (1992), Zheng (1996), and Hong and White (1995)..

[2] In general, essentially any real analytic function is generically comprehensive, including the exponential, logistic and sin + cos: see Stinchcombe and White (1997).

does not influence the process under the null, hence it must be treated as a nuisance parameter. Teräsvirta (1994) and many others suggest performing the polynomial regression tests for various delay values, say d, and selecting that d which generates the lowest test p-value. This is mathematically equivalent to generating an LM sup-statistic over possible d-values, a statistic known to have a non-standard limit distribution. Nevertheless, in the literature the standard practice is simply to employ p-values derived from the chi-squared distribution.

Here, we abstract from asymptotic theory and focus entirely on small sample performance of a class of tests universally over-looked in the above smooth transition literature. In a broad simulation study of linear AR, STAR and bilinear processes, we employ Hansen's (1996) method for approximating the null distribution of an LM sup-statistic based on the sample null score, and demonstrate the superior strengths of the resulting hybrid test method. Of separate interest, our simulation study also provides a rare glimpse into the comparative strengths of Bierens' (1990) and Hansen's (1996) competing solutions to the dilemma of asymptotic non-standard null distributions: we demonstrate that while a Bierens-type test for model specification against STAR alternatives provides optimal power, Hansen's (1996) p-value method is the best technique for analyzing the test statistic's distribution for arbitrary sample sizes. Indeed, the combined LM sup-test with bootstrapped p-values typically dominates the popularly espoused polynomial regression test for STAR processes, in particular for logistic STAR processes, and in general for STAR processes "far" from linear. Moreover, in many cases the STAR test dominates the Bierens test and the popularly employed neural test of "neglected nonlinearity", cf. Lee *et al* (1993), for detection of general model mis-specification.

Of particular note, our simulation study is substantially less restrictive than previous such studies: we do not fix *any* STAR parameters, and therefore control for the fact that *a priori* chosen parameters may bias test results (e.g. Luukkonen *et al*, 1988, Teräsvirta, 1994, Skalin, 1998). We find over a broad range of admissible STAR parameter values that a sup-LM statistic out-performs extant tests of STAR and general nonlinearity.

Finally, Skalin (1998) considers a Likelihood Ratio test of STAR nonlinearity in a simulation based environment, and employs Hansen's (1996) bootstrap method for approximating the asymptotic p-value. The author finds the polynomial regression method dominates the bootstrapped LR statistic. Contrary to a score test, an LR test requires estimation of a specific alternative model. Even when using efficient maximum likelihood estimation, STAR models are renowned for their difficulty to provide sharp estimates of imperative transition function coefficients. Typically the imperative scale parameter γ estimate, on which the LR test hinges, appears to be insignificant, even in controlled experiments where the data generating structure is STAR: see Teräsvirta (1994) and Franses and van Dijk (2000). Thus, the power of such a test should be held suspect, and it is not surprising, therefore, that Skalin (1998) finds the LR test to be dominated by the provably non-consistent polynomial F-test method. Moreover, apparently the only consistent parametric tests of functional form belong to the class of conditional moment tests developed in Bierens (1982, 1990) and Bierens and Ploberger (1997) which lends itself specifically to an LM test framework. Because it is this class of tests that can be employed for a consistent test of linear autoregression against smooth transition alternative, we ignore LR tests.

The rest of this paper contains the following topics. In Section 2 the sup-LM test is outlined, and the simulation study is performed in Section 3. Tables can be found in the appendix.

2. Lagrange Multiplier Sup-Test Consider a standard two-regime STAR model:

$$y_t = \phi_1' x_{1t} + \phi_2' x_{2t} F_t(d, \gamma, c) + u_t. \tag{3}$$

For simplicity of notation, assume regressors are identical across regimes, x_{1t} = x_{2t} = x_t, a $k \times 1$ vector. The fundamental null hypothesis of linearity maintained throughout states

$$H_0 : \phi_2 = 0. \tag{4}$$

Under the null hypothesis, therefore,

$$y_t = \phi_1' x_t + \epsilon_t, \tag{5}$$

where $\epsilon_t = u_t$. Under the null, the transition parameters γ, c, and d are unspecified: the hypothesis holds for any values, and therefore we treat the transition parameters as nuisance parameters.

Alternatively, when $\gamma = 0$, the transition function is a constant (0 for the exponential, and 1/2 for the logistic) and the STAR model collapses to a linear AR model, for any value of ϕ_2, c, and d. The hypothesis $H_0 : \gamma = 0$ is the favored focal point in the STAR test literature. Here, we focus on (8), and develop an associated LM statistic.

Denote by θ the vector of all parameters $(\phi_1', \phi_2', d, \gamma, c)'$. For compactness, define the sub-vector $\varphi = (d, \gamma, c)'$. It is straightforward to show that the null score obtains the representation

$$s_n(\theta)|_{H_0} = s_n(0, \varphi) = \left[\begin{array}{c} n^{-1} \sum_{t=1}^{n} \epsilon_t x_t \\ n^{-1} \sum_{t=1}^{n} \epsilon_t x_t F_t(\varphi) \end{array} \right]. \tag{6}$$

Using least squares estimates from the null model, we obtain

$$\begin{aligned} \hat{s}_n(0, \varphi) &= n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t x_t F_t(\varphi) \qquad (7) \\ &= n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t z_t \\ \hat{\epsilon}_t &= y_t - \hat{\phi}_1' x_t, \quad z_t = x_t F_t(\varphi). \end{aligned}$$

The LM statistic, therefore, satisfies

$$T_n(\varphi) = n \hat{s}_n(0, \varphi)' \hat{V}(\varphi)^{-1} \hat{s}_n(0, \varphi) \tag{8}$$

where standard asymptotic algebra shows (see, e.g., Bierens, 1990)

$$\begin{aligned}\hat{V}(\varphi) &= \frac{1}{n}\sum\nolimits_{t=1}^{n} \hat{\epsilon}_t^2 \left[F_t(\varphi)I_k - \hat{b}(\varphi)\hat{A}^{-1}\right]' x_t x_t' \left[F_t(\varphi)I_k - \hat{A}^{-1}\hat{b}(\varphi)\right] \quad (9)\\ \hat{b}(\varphi) &= \frac{1}{n}\sum\nolimits_{t=1}^{n} F_t(\varphi)x_t x_t', \quad \hat{A} = \frac{1}{n}\sum\nolimits_{t=1}^{n} x_t x_t',\end{aligned}$$

and I_k denotes the k-dimensional identity matrix.

Finally, define the sup-statistic,

$$g_n = \sup\nolimits_{\varphi} T_n(\varphi), \tag{10}$$

where the supremum is taken over feasible values of the coefficient vector, $\varphi = (d, \gamma, c)'$: see below for details.

3. Simulation Study We now investigate the empirical size and power properties of the STAR sup-statistic $g_n = \sup T_n(\varphi)$. Our simulations are based on the following models:

$$\begin{aligned}&H_0 : y_t = \phi_1' x_t + \epsilon_t\\ &H_1^L : y_t = \phi_1' x_t + \phi_2' x_t \left(1 + \exp[-\gamma(y_{t-d} - c)]\right)^{-1} + \epsilon_t\\ &H_1^E : y_t = \phi_1' x_t + \phi_2' x_t \exp[-\gamma(y_{t-d} - c)^2] + \epsilon_t\\ &H_1^{BL} : y_t = \phi_1' x_t + y_{t-1}\epsilon_{t-1} + \epsilon_t\end{aligned}$$

where ϵ_t is *iid* standard normal[3], and $x_t = (1, y_{t-1}, ..., y_{t-p})'$ for some $p > 0$. Under H_0 the true data generating process is linear; under H_1^L and H_1^E the true process is a 2-regime LSTAR and ESTAR, respectively; under H_1^{BL} the process is a hybrid bilinear-autoregression.

3.1 Set-up We consider sample sizes $n = 100$, 500, and 1000: in each case, we generate $3n$ observations, and retain the last n observations in order to reduce dependence on starting values. For each simulated series, the order p is randomly chosen from the set $\{1, ..., 10\}$, and ϕ is randomly chosen from the uniform hypercube $[-1.5, 1.5]^{p+1}$. Moreover, the scale parameter γ is randomly selected from the uniform interval $[.05, 5]$, the threshold c is randomly selected from $[-.5, .5]$, and delay d is randomly chosen from the integer set $\{1, ..., p\}$. Because typical asymptotic considerations require the null model to be covariance stationary, only vectors ϕ_1 with characteristic polynomial roots outside the unit circle are considered.

We generate 1000 replications of each series above. For each series, a linear model is estimated and the resulting residuals are tested at the 5%-level. In order to specify the null model, we employ a minimum AIC model selection criterion for the order p over the integer set $\{1, ..., 10\}$.

[3] All simulations are performed using the GAUSS 5.0 software. Code is available upon request.

3.2 Tests In order to test for linearity, consider model (3). The STAR sup-LM test is based on the score weight $z_t = x_t F_t(\varphi)$, cf. (7). The test is performed based on grid-searches over d, γ, and c. Following the standard rule of thumb (see, e.g., Teräsvirta, 1994), possible threshold values c are limited to the interval between the lower and upper 15^{th}-quantiles of y_t, denoted $y_{[.15]}$ and $y_{[.85]}$. Moreover, for γ we search over the interval $[.1, 10]$, and the candidate delay values are restricted to the interval set $\{1, 2, 3\}$. Although the simulated STAR series use a randomly selected scale γ with lower bound .05, the tests themselves must use a larger lower bound (.10) due to covariance matrix singularity for small samples: the weight $x_t F_t(\varphi) \approx x_t \times \zeta$ for some scalar-constant $\zeta > 0$ when γ is close to zero, in which case the asymptotic covariance matrix estimator $\hat{V}(\varphi)$ becomes singular due to machine error. We use both Bierens' (1990) criterion technique for generating an asymptotic χ^2-statistic, and Hansen's (1996) simulated p-value method in other to approximate the true null distribution of g_n.

Bierens' (1990) criterion technique for generating an asymptotically χ^2-statistic is performed as follows. Let $\varphi^* = \arg\max_\varphi T_n(\varphi)$, and let $\tilde{\varphi}$ denote a nuisance vector randomly selected independent of the sample of data. Define the respective LM statistics $T_n(\varphi^*)$ and $T_n(\tilde{\varphi})$. For arbitrary parameters $\psi > 0$ and $\rho \in (0, 1)$, we select $\hat{\varphi}$ such that

$$\begin{aligned} \hat{\varphi} &= \tilde{\varphi} \ \text{ if } T_n(\varphi^*) - T_n(\tilde{\varphi}) \leq \psi n^\rho \\ \hat{\varphi} &= \varphi^* \ \text{ if } T_n(\varphi^*) - T_n(\tilde{\varphi}) > \psi n^\rho. \end{aligned}$$

Under H_0 and appropriate assumptions governing dependence, the resulting statistic $T_n(\hat{\varphi})$ converges in law to a χ^2-random variable with k-degrees of freedom: see Bierens (1990: Theorems 4-5). In lieu of simulation evidence reported in Bierens (1990), we use $\psi = .25$ and $\rho = .5$. This is the *STAR_Bier* test.

Hansen's (1996) method involves simulating the null distribution of the sup-score $\hat{s}(0, \varphi^*)$, where $\varphi^* = \arg\max_\varphi T_n(\varphi)$. We draw $n \times J$ *iid* random variables $u_{t,j} \sim N(0, 1)$, $t = 1...n$, $j = 1...J$, and generate a sample of J-scores and J-test statistics:

$$\begin{aligned} \hat{s}_{n,j}(0, \varphi^*) &= \frac{1}{n} \sum\nolimits_{t=1}^{n} \hat{\epsilon}_t x_t F_t(\varphi^*) u_{t,j} \\ T_{n,j}(\varphi^*) &= n \hat{s}_{n,j}(0, \varphi^*)' \hat{V}(\varphi^*)^{-1} \hat{s}_{n,j}(0, \varphi^*). \end{aligned} \tag{11}$$

The approximate p-value of the sup-statistic g_n is simply the frequency with which $T_{n,j}(\varphi^*) > g_n$ occurs. For all simulations, we set $J = 500$. This is the *STAR_han* test.

We also perform the neural test of neglected nonlinearity (Lee *et al,* 1996), the Bierens test, the McLeod-Li test, the RESET test, and the polynomial regression test of Luukkonen *et al* (1988) and Teräsvirta (1994).

The Bierens test is similar to the STAR test, except the scalar weight $z_t = F_t(\varphi)$ is used. In this sense, we can simply interpret the Bierens test as a test against a restricted STAR process with second regime slopes equal to zero: see also Franses and van Dijk (2000) on a related point. For the Bierens test, we employ both Bierens' (1990) criterion, with $\psi = .5$ and $\rho = .25$ (*BIER*); and we use Hansen's (1996) method for evaluating the true distribution of the Bierens sup-statistic (*BIER_han*). In this manner, we control for the possibility that differences between the STAR sup-test and all other tests is merely due to the use of Hansen's (1996) method.

The neural test is equivalent to the Bierens test (i.e. $z_t = F_t(\varphi)$), except all nuisance parameters d, γ, and c are randomly selected from their respective intervals, detailed above.

For the standard STAR polynomial test, we estimate models of the form

$$y_t = \phi_1' x_{1t} + \sum\nolimits_{i=1}^{L} \beta_i' \tilde{x}_t y_{t-1}^i + u_t, \tag{12}$$

where $\tilde{x}_t = (y_{t-1}, ..., y_{t-p})$. Under a null of linearity against an LSTAR alternative, $L = 3$ and (12) implies $\beta_i = 0$, $i = 1..3$. Under a null of linearity against an ESTAR alternative, $L = 4$ and (12) implies $\beta_i = 0$, $i = 1..4$. In order to decide between LSTAR and ESTAR alternatives based on the polynomial regression, Teräsvirta (1994) suggests a test of $H_0 : \beta_i = 0$, $i = 1..4$ first in order to substantiate concern for STAR nonlinearity at all, then a sequence of F-tests on parameter sub-sets from (12). Because we are interested in whether the test procedure can find *any* deviation from the null of linearity, we do not pursue the test sequence approach and simply report null rejection frequencies based on tests of (12) with $L = 3$ or 4. Rejection in either case is argued to be consistent with evidence in favor of STAR nonlinearity, cf. Luukkonen *et al* (1988). These are the *POLY_L* and *POLY_E* tests, respectively.

For the McLeod-Li test, we perform a standard portmanteau test on the squared null residuals for lags 1...3. For the RESET test, we follow the procedure detailed in Thursby and Schmidt (1977) by estimating the auxiliary regression based on the null residuals $\hat{u}_t$,

$$\hat{u}_t = \beta_0' x_t + \sum\nolimits_{i=2}^{L} \sum\nolimits_{j=2}^{k} \beta_{i,j} x_{t,j}^i + w_t, \tag{13}$$

where we set $L = 3$. A standard LM test for the linearity hypothesis $H_0 : \beta_{i,j} = 0$ is performed.

For all LM tests employed in this study, covariance matrix estimators robust to unknown forms of conditional heteroscedasticity are used.

3.3 Results Test results for H_0 are contained in Table 1, and Table 2 contains empirical powers for LSTAR, ESTAR and bilinear alternatives. For linear processes, the STAR tests compare well with the popularly used neural and polynomial regression tests. The polynomial and STAR tests tend to under-reject the null, while only the neural and exponential Bierens tests closely approach the 5% level.

Under an LSTAR alternative, the STAR tests evaluated by bootstrapped p-values dominate all other tests for sample sizes of 100 and 500. For large $n =$ 1000 the STAR test only slightly out-performs the polynomial regression tests, yet still dominates all other tests. Under an ESTAR alternative, by comparison, the STAR tests do not perform as favorably as the polynomial regression test, however all tests, except the Bierens test, generate relatively low empirical rejection frequencies. The Bierens test with an exponential weight is more adept at detecting linear model mis-specification in the presence of true ESTAR nonlinearity than the polynomial regression test. This also suggests in practice that a neural network model may be found to improve model performance even when the traditional STAR test suggests a linear model in adequate.

In general, the use of Bierens' (1990) criterion method is dramatically suboptimal for the STAR sup-test, but not the Bierens test. Of course, the problem may simply be due to the fixed criterion parameters values for ψ and ρ. In any event, our simulations strongly suggest Hansen's (1996) bootstrapped p-value method for the STAR test renders a highly competitive test method, however does not improve the performance of the Bierens test. Because bootstrapped p-values are increasingly de riguer in practice, and because they impressively aid STAR test performance here, we do not consider other values of ψ and ρ.

Moreover, the Bierens test, even when analyzed by bootstrapped p-values, is substantially out-performed by the STAR and neural tests for tests on bilinear processes, the class of processes used in the simulation study of Bierens (1990). The McLeod-Li test, however, has optimal power against Gaussian bilinear processes (see McLeod and Li, 1983), hence it is not surprising how well the test performs in our study. For STAR nonlinearity, however, the McLeod-Li test is outperformed by the Bierens, neural and in particular the STAR tests.

Recall that we randomize the transition scale $\gamma \in [.05, 5]$ for STAR processes. Values closer to zero imply a STAR model with very slow regime transition such that the process appears to be "nearly linear". Thus, our simulation study demonstrates that over a broad spectrum of transition velocities the sup-LM test dominates *on average* the polynomial regression method, in particular against LSTAR alternatives. However, the sup-LM statistic used in the present study is shown by Andrews and Ploberger (1994) to be the limit of an optimal test (admissible for any alternative, hence any non-zero value of ϕ_2 and γ), and effectively directs power toward distant deviations from the null. Thus, despite being asymptotically consistent against any deviation from the null, the statistic is expected to provide more power against deviations "far" from linearity (i.e. large γ) for small samples.

We check this by performing an identical simulation study of only STAR processes with a fixed $\gamma = 3$. All other simulation specifications remain as above. Results are contained in Table 3. For STAR process that are more "distant" from a linear autoregression, the sup-LM test demonstrates its comparative power lift both relative to itself when performed on STAR processes that may be "nearly linear", and relative to the polynomial regression method. Using bootstrapped p-values and for $n \geq 500$, the STAR test with either exponential of logistic weights $x_t F_t(\varphi)$ correctly detects LSTAR nonlinearity in over 77% of such series, while the polynomial test correctly rejects the null in favor of STAR in 43.50% (58.40%) series with sample size 500 (1000). Of particular note, the STAR tests improved on empirical power for small $n = 100$ by nearly

30% relative to simulations with randomized γ, while the polynomial F-test results in about 5% more power, admittedly an increase of over 100% relative to the previous rejection rate with a randomized scale γ. However, even the RESET test performs better than the standard polynomial STAR test for "strong" STAR processes. In general, using either an exponential or logistic weight the STAR sup-LM test dominates all tests when the true process is LSTAR.

Not only does the STAR test provide power leverage against unknown arbitrary deviations from the null (e.g. possibly small γ and bilinear nonlinearity), but against strong forms of smooth transition nonlinearity (i.e. large γ) and in particular against logistic STAR nonlinearity. The present simulation study persuasively demonstrates that the universally employed polynomial STAR test is sub-optimal relative to a sup-LM test, as well as the consistent Bierens and neural tests. Indeed, the polynomial test may not even be an appropriate test methodology for detecting which STAR form, exponential or logistic, best approximates the true data generating process. We may well argue that the best test approach is a consistent STAR test of linear autoregression against either STAR form, where the subsequent decision between STAR forms is informed by economic theory and policy considerations.

Appendix

Table 1

$H_0 : AR(p)$			
n	100	500	1000
STAR_Han_L	.0330[a]	.0040	.0010
STAR_Han_E	.0360	.0070	.0030
STAR_Bier_L	.0580	.0140	.0060
STAR_Bier_E	.0840	.0340	.0130
NEURAL_L	.0390	.0460	.0440
NEURAL_E	.0390	.0590	.0390
BIER_han_L	.0110	.0240	.0180
BIER_han_E	.0400	.0580	.0430
BIER_L	.0170	.0260	.0150
BIER_E	.0420	.0470	.0490
POLY_L	.0010	.0030	.0020
POLY_E	.0010	.0030	.0020
RESET	.0450	.0380	.0490
ML-1[b]	.0520	.0700	.0860
ML-2	.0570	.0880	.0910
ML-3	.0640	.1050	.0950

Notes: a. p-values less than .00005 are reported as .0000;
b. ML-h denotes the ML test with h-lags.

Table 2

n	100			500			1000		
H_1 : $STAR$	H_1^L	H_1^E	H_1^{BL}	H_1^L	H_1^E	H_1^{BL}	H_1^L	H_1^E	H_1^{BL}
STAR_Han_L	.2790	.0700	.1210	.5500	.1450	.3400	.6650	.2400	.5050
STAR_Han_E	.2830	.0720	.1490	.5660	.1650	.4100	.6840	.2900	.5750
STAR_Bier_L	.1520	.0690	.1040	.0970	.0320	.1310	.0840	.0280	.1740
STAR_Bier_E	.2000	.1050	.1420	.1670	.0440	.1770	.1330	.0420	.2190
NEURAL_L	.1830	.0710	.2110	.4290	.1950	.4690	.5020	.2880	.5760
NEURAL_E	.1720	.0790	.2060	.3930	.1940	.4790	.4770	.2720	.5630
BIER_han_L	.1650	.0320	.0290	.4870	.1300	.0470	.5810	.2170	.0710
BIER_han_E	.2350	.1120	.2100	.5310	.3210	.2410	.6180	.4360	.2360
BIER_L	.2350	.0500	.0970	.5070	.1590	.0720	.5970	.2350	.0810
BIER_E	.1720	.1130	.2210	.5800	.3450	.2350	.6220	.4670	.2410
POLY_L	.0430	.0040	.0180	.4330	.1800	.0290	.6510	.3820	.0180
POLY_E	.0430	.0040	.0180	.4330	.1800	.0290	.6510	.3820	.0180
RESET	.2610	.0210	.0920	.4110	.1090	.0060	.5150	.2710	.0010
ML-1	.1130	.0310	.5160	.3060	.0650	.9750	.3610	.1020	.9920
ML-2	.1240	.0500	.5170	.3660	.1270	.9870	.4050	.1690	.9990
ML-3	.1510	.0780	.5240	.4090	.1510	.9960	.4970	.2120	.9990

Table 3

n	100		500		1000	
$H_1 : STAR$, $\gamma = 3$	H_1^L	H_1^E	H_1^L	H_1^E	H_1^L	H_1^E
STAR_Han_L	.5670	.1020	.7700	.3150	.7910	.4660
STAR_Han_E	.5600	.1480	.8020	.3660	.8070	.5300
STAR_Bier_L	.1760	.0830	.1220	.0660	.1060	.0430
STAR_Bier_E	.2250	.1410	.1590	.1060	.1280	.0710
NEURAL_L	.3860	.2150	.7260	.5090	.7680	.5720
NEURAL_E	.3440	.2050	.6770	.4920	.7090	.5700
BIER_han_L	.3820	.0710	.6990	.2530	.7430	.4040
BIER_han_E	.4740	.3340	.7570	.5950	.7970	.7170
BIER_L	.4060	.1980	.6460	.4370	.7160	.5560
BIER_E	.4580	.3670	.7040	.6540	.7700	.7350
POLY_L	.0910	.0070	.4350	.2290	.5840	.4860
POLY_E	.0910	.0070	.4350	.2290	.5840	.4860
RESET	.3850	.0410	.6320	.2030	.6990	.3680
ML-1	.1270	.0490	.3250	.0890	.3820	.1190
ML-2	.1490	.0710	.4010	.1440	.4490	.1850
ML-3	.1780	.0850	.4310	.1690	.5160	.2280

References

[1] Bierens, H. J., 1982, Consistent Model Specification Tests, Journal of Econometrics 20, 105-134.

[2] Bierens, H. J., 1990, A Consistent Conditional Moment Test of Functional Form, Econometrica 58, 1443-1458.

[3] Bierens, H. J. and W. Ploberger, 1997, Asymptotic Theory of Integrated Conditional Moment Tests, Econometrica 65, 1129-1151.

[4] Chan, K.S. and H. Tong, 1986b, On Tests for Nonlinearity in Time Series Analysis, Journal of Time Series Analysis 5, 217-228.

[5] Davies, R.B., 1987, Hypothesis Testing When a Nuisance Parameter is Present Only under the Alternative, Biometrika 74, 33-43.

[6] van Dijk, D., T. Teräsvirta and P.H. Franses, 2000, Smooth Transition Autoregressive Models-A Survey of Recent Developments, Econometric Institute Research Report EI2000-23/A, Erasmus University.

[7] Escribano, A. and O. Jordá, 1999, Improved Testing and Specification of Smooth Transition Autoregressive Models, in P. Rothman (ed.), Nonlinear Time Series Analysis of Economic and Financial Data, pp. 289-319 (Boston: Kluver).

[8] Gonzalez-Rivera, G., 1998, Smooth-transition GARCH models, Studies in Nonlinear Dynamics and Econometrics 3, 61-78.

[9] Hagerud, G.E., 1997, A New Nonlinear GARCH Model, Unpublished Ph.D. thesis, IFE, Stockholm School of Economics.

[10] Hansen, B., 1996, Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis, Econometrica 64, 413-430.

[11] Holley, A., 1982, A Remark on Hausman's Specification Test, Econometrica 50, 749-760.

[12] Hong, Y. and H. White, 1995, Consistent Specification Testing via Nonparametric Series Regression, Econometrica 63, 1133-1159.

[13] Keenan, D.M., 1985, A Tukey Nonadditivity-Type Test for Time Series Nonlinearity, Biometrika 72, 39-44.

[14] Kuan, C. and H. White, 1994, Artificial Neural Networks: An Economic Perspective, Econometric Reviews 13, 1-91.

[15] Lee, T., 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.

[16] Luukkonen, R., P. Saikkonen and T. Teräsvirta, 1988, Testing Linearity against Smooth Transition Autoregressive Models, Biometrika 75, 491-9.

[17] Madieros, M.C. and A. Veiga, 2000, Diagnostic Checking in a Flexible Nonlinear Time Series Model, manuscript, Dept. of Electrical Engineering, Catholic University of Rio de Janeiro.

[18] Ramsey, J.B., and P. Schmidt, 1976, Some Further Results on the Use of OLS and BLUE Residuals in Specification Error Tests, Journal of the American Statistical Association 71, 389-390.

[19] Saikkonen, P. and R. Luukkonen, 1988, Lagrange Multiplies Tests for Testing Nonlinearities in Time Series Models, Scandinavian Journal of Statistics 15, 55-68.

[20] Skalin, J., 1998, Testing Linearity against Smooth Transition Autoregression Using A Parametric Bootstrap, Working Paper Series in Economics and Finance No. 276, Dept. of Economic Statistics, Stockholm School of Economics.

[21] Stinchcombe, M.B. and H. White, 1998, Consistent Specification Testing with Nuisance Parameters Present Only Under the Alternative, Econometric Theory 14, 295-325.

[22] Teräsvirta, T., 1994, Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models, Journal of the American Statistical Association 89, 208-218.

[23] Thursby, J., and P. Schmidt, 1977, Some Properties of Tests for Specification Error in a Linear Regression Model, Journal of the American Statistical Association 72, 635-641.

[24] Tong, H., 1983, Threshold Models in Non-Linear Time Series Analysis (New York: Springer-Verlag).

[25] Wooldridge, J., 1991, On the Application of Robust, Regression- Based Diagnostics to Models of Conditional Means and Conditional Variances, Journal of Econometrics 47, 5-46.

[26] Yatchew, A.J., 1992, Nonparametric Regression Tests Based on Least Squares, Econometric Theory 8, 435-451.

[27] Zheng, J., 1996, A Consistent Test of Functional Form via Nonparametric Estimation Techniques, Journal of Econometrics 75, 263-289.