An ARCH Robust STAR Test

Michael K. Andersson, Bruno Eklund and Johan Lyhagen

Working Paper Series in Economics and Finance No. 317
Stockholm School of Economics
May 1999

Abstract

The LM type linearity test for STAR nonlinearities is severely distorted when the process is governed by conditional heteroskedasticity. In order to correct the test we propose a parametric bootstrap. It is shown, by means of Monte Carlo methods, that the bootstrap test is almost exact.
Key words: Smooth transition autoregressive models; Bootstrap; Parametric resampling; Size distortion; Power
JEL Classification: C12; C22

1 Introduction

A nonlinear behavior of an economic variable is intuitively appealing. For example, it is plausible that unemployment rates react differently in periods of contraction and expansion, due to labor market regulations and employment costs, such as training and advertisement. Similarly, business cycles may also be argued to exhibit nonlinear features.

The first step in a nonlinear modeling procedure is to test the hypothesis of linearity. Due to the complexity of the nonlinear world, a specific type of nonlinearity generally constitutes the alternernative hypothesis. One alternative model family is smooth transition autoregressions (STAR), which may be regarded as two state autoregressive models with smooth transitions between the states. The modeling procedure is described by Teräsvirta (1994a).

When testing for a STAR type nonlinearity it is important that the second moment is correctly specified. For instance, Lundberg and Teräsvirta (1998) demonstrate that conditional heteroskedasticity distorts the test and that a robustified version of the test statistic suffers from a severe power loss. As an alternative, we introduce two bootstrap procedures in order to adjust the critical values. The bootstrap testing approach is expected to be close to exact, both for independent and heteroskedastic errors, and still have power against STAR alternatives.

It is demonstrated that the procedure works well if the parametric approximation of the second moment structure is reasonably close to the true.

The paper is outlined as follows. Section 2 provides a brief introduction to STAR models and the linearity test, while Section 3 presents the bootstrap test. A Monte Carlo simulation in Section 4 investigates the properties of the tests and a concluding section ends the paper.

2 Smooth Transition Autoregression Models

A pth-order smooth transition autoregressive (STAR) model has the general form

$$y_t = \theta_0 + \theta' w_t + (\phi_0 + \phi' w_t) F(\gamma, y_{t-d}, c) + a_t, \tag{1}$$

where a_t is an iid error process, $E(a_t) = 0$, $E(w_t a_t) = 0$, $Var(a_t) = \sigma^2$, and where $w_t = (y_{t-1}, \ldots, y_{t-p})'$ enters the model with constant parameter vectors $\theta' = (\theta_1, \theta_2, \ldots, \theta_p)'$ and $\phi' = \left(\phi_1, \phi_2, \ldots, \phi_p\right)'$. The nonlinearity is introduced by the transition function $F(\cdot)$. Two common parametrizations of F are considered, namely the exponential, yielding an ESTAR model,

$$F(\gamma, y_{t-d}, c) = 1 - \exp\left\{-\gamma (y_{t-d} - c)^2\right\} \ , \gamma > 0 \tag{2}$$

and the logistic, LSTAR,

$$F(\gamma, y_{t-d}, c) = \frac{1}{1 + \exp(-\gamma(y_{t-d} - c))} \ , \gamma > 0. \tag{3}$$

The fact that the model is linear when $F = 0$ is utilized when performing the linearity test. Under the null hypothesis, the problem with unidentified parameters are circumvented using a three-term Taylor expansion of F around $\gamma = 0$, implying that the test can be performed via the auxiliary regression

$$y_t = \beta_0 + \beta' w_t + \delta_1' w_t y_{t-d} + \delta_2' w_t y_{t-d}^2 + \delta_3' w_t y_{t-d}^3 + u_t. \tag{4}$$

The linearity test, proposed by Teräsvirta (1994b), corresponds to testing $H_0 : \delta_1 = \delta_2 = \delta_3 = 0$ in (4), and has power against both LSTAR and ESTAR alternatives. Thus, the test can not only detect nonlinearity in the data, but also suggest the type of nonlinearity, see Teräsvirta (1994a,b).

3 The Bootstrap Testing Procedure

The bootstrap, described by Efron and Tibshirani (1993), provides a feasible method for estimation of the small-sample distribution of a statistic. The basic principle is to approximate this distribution by a bootstrap distribution, which is retrieved by simulation. If the exercise is bootstrap hypothesis testing the bootstrap samples must obey the null hypothesis and, as far as possible, resemble the real sample.

3.1 The Bootstrap Test

Asymptotic theory is only exact if the p-value is independent of the actual data generating process and sample size, which is usually not the case. A small sample solution is to replace the p-value by the bootstrap counterpart, which can be estimated as

$$\hat{p}^* \left(\hat{f} \right) = R^{-1} \sum_{r=1}^{R} I \left(f_r{}^* > \hat{f} \right), \tag{5}$$

where R is the number of bootstrap replicates, $I(\cdot)$ the usual zero/one indicator function, $\hat{f}$ a realized value of the test statistic f based on a sample $\mathbf{y} = \{y_1, ..., y_T\}$ and $f_r{}^*$ the value of the same test statistic, based on the bootstrap sample $\mathbf{y}_r^* = \{y_{r1}^*, ..., y_{rT}^*\}$.

The theory of bootstrap testing is developed by Davidson and MacKinnon (1996, 1999a). It is shown that if the test statistic is (asymptotically) pivotal, that is independent of nuisance parameters, the size-distortion refinement is of order $T^{-1/2}$ when using the bootstrap p-value compared to the corresponding asymptotic. A further refinement, usually also of order $T^{-1/2}$ is obtained whenever the test statistic is independent of the bootstrap

DGP. Moreover, the power of a bootstrap test, based on a pivotal statistic, is generally close to the size-adjusted asymptotic test. Even if the statistic is only close to pivotal this is true in most cases.

To achieve an exact test the number of bootstrap replicates must be chosen such that $\alpha (R+1)$, where α is the desired level, is an integer. Davidson and MacKinnon (1999*b*) demonstrate that a small number of bootstrap replicates imply a loss in power and that at least $R = 399$ is required to guarantee a power loss of no more than 1% at the 0.05 level. The size of a bootstrap test is less sensitive to the number of replicates.

3.2 Construction of the Bootstrap Samples

For the construction of the bootstrap samples we use a model-based approach, which is natural since a well-defined model constitutes the null hypothesis. The bootstrap B_A, which is constructed to preserve ARCH(1) dependence in the residuals, is conducted as follows:

1. Estimate the AR(p)-ARCH(1) model

$$\begin{aligned} \left(1 - \phi_1 B - ... - \phi_p B^p\right)(x_t - \mu) &= a_t, \quad a_t \,|\mathcal{I}_{t-1} \sim N(0, \omega_t) \\ \omega_t &= \beta_0 + \beta_1 a_{t-1}^2 \end{aligned} \qquad (6)$$

   which clearly obeys the null-hypothesis. The autoregressive order p is determined by the Akaike information criterion (AIC) and the parameters are estimated through maximization of the log-likelihood function.

2. Due to the assumed normality of the disturbances a_t in (1), the bootstrap residuals $\{a_t^*\}$ are constructed accordingly; let ε_t^* be an independent draw from a $N(0,1)$ distribution, then the bootstrap residuals are computed as

$$\begin{aligned} \hat{\omega}_t &= \hat{\beta}_0 + \hat{\beta}_1 a_{t-1}^{*2} \\ a_t^* &= \varepsilon_t^* \sqrt{\hat{\omega}_t} \end{aligned}$$

3. The bootstrap samples $\mathbf{x}_r^*$, $r = 1, ..., 399$, are created recursively by the equation

$$x_{r,t}^* = \hat{\mu} + \hat{\phi}(B)^{-1} a_t^*,$$

   where $\hat{\phi}(B)$ is the estimated polynomial of (6).

Of course, the procedure is not limited to ARCH(1) errors, it can easily be extended to GARCH processes and the lag-orders may be selected from data, e.g. by some information criterion. The reason for only considering ARCH processes and a pre-specified lag-order is purely time saving, bootstrap-Monte Carlo studies are computationally demanding.

For the sake of comparison, a simple bootstrap version, which ignores the ARCH, is also included. This resampling, denoted B_S, draws residuals a_t^* independently direct from a normal distribution with mean zero and variance $s_{\hat{a}}^2$.

4 Properties of the Tests

The properties, that is the size and power, of the tests are investigated by means of a Monte Carlo simulation. The study involves 10 000 replicates (series), where each series is tested against STAR using the asymptotic test and the different bootstrap tests. The rejection frequencies of the nonlinear null hypothesis are calculated and evaluated.

4.1 Size

The empirical size of the tests are computed for first order autoregressions

$$y_t = \phi y_{t-1} + a_t, \quad t = 1, ..., T, \tag{7}$$

where the serial length T is set to 50, 100 and 200, and the parameter ϕ is selected as $\{0.0, \pm 0.2, \pm 0.6, \pm 0.9\}$. To reduce the initial value effect an additional 100 observations are generated and discarded. The errors are constructed to display independence or conditional heteroskedasticity of ARCH(1) type. In the independent case, a_t is normally distributed with zero mean and unit variance. The ARCH disturbances are conditionally distributed as $a_{t|t-1} \sim N(0, \omega_t)$, where $\omega_t = \sigma_a^2(1-\beta) + \beta a_{t-1}^2$ and $\beta < 1$. If σ_a^2 is set to unity, the parametrization implies that the unconditional variance of a_t equals unity, and the parameter β is selected as 0.5 and 0.9.

Table 1 presents the sensitivity, at a nominal five percent level of significance, of the empirical size with respect to AR parameters when the true disturbance process is normally distributed. The estimated size is close to the nominal for the asymptotic test already at sample size 50. Moreover, the simple bootstrap test B_S is at least as accurate as the asymptotic test and the routine B_A, which accounts for (non-existing) ARCH effects, is somewhat conservative. However, as the serial length grows the size approaches the desired five percent.

The estimated sizes when the error term is conditionally heteroskedastic are reported in Table 2. The asymptotic test is seriously distorted when the ARCH effect is governed by $\beta = 0.5$, and the empirical size diverge from the nominal. This behavior is inherited by the simple bootstrap test. When allowing for ARCH in the resamples, the bootstrap test works very well in the sense that the empirical size is close to the nominal. All tests react similarly for generated $\beta = 0.9$ processes. However the distortion of the asymptotic and simple bootstrap tests are enhanced and at $T = 200$ they peak at almost 80%. The ARCH bootstrap test is slightly over-sized at $T = 50$, $6 < \hat{\alpha} < 7.10$, but the distortion is corrected for in larger samples. For the largest sample the test is close to exact. The results demonstrate that a correctly specified bootstrap testing procedure works extremely well, in the sense that its empirical size is close to the nominal.

4.2 Power

The experiment examining the power of the tests covers the specification

$$y_t = 1.8y_{t-1} - 1.06y_{t-2} + (0.02 - 0.9y_{t-1} + 0.795y_{t-2})\,F(\gamma, 0.02, y_{t-1}) + a_t, \tag{8}$$

where the transition function assumes the forms (2) and (3). As Teräsvirta (1994), from whom the specifications are borrowed, we set the γ to 20 or 100 for the ESTAR, and to 100 or 1 000 for the LSTAR model, and the members of $\{a_t\}$ are generated as independently and normally distributed with mean zero and variance 0.0004. Moreover, we also generate disturbances that are conditionally heteroskedastic, as above, with an unconditional variance of 0.0004. The empirical rejection frequencies, that is the estimated power, is studied for series of length $T = 50, 100, 200$ and 400.

The rejection percentage, i.e. the empirical power of the tests, for generated STAR processes are reported in Figure 1. Given uncorrelated errors, the asymptotic test has high estimated power already at small sample sizes. For the LSTAR specification, the rejection frequencies increase with γ, whereas the opposite is found for ESTAR processes. This is anticipated considering the shape of the transition functions. The empirical power of the simple bootstrap test is close to the asymptotic test, and the ARCH bootstrap exhibit a (often marginally) lower power. This is explained by the conservative behaviour of the procedure.

When the disturbances are conditionally heteroskedastic we only examine the power of the ARCH bootstrap, since the other tests are seriously distorted. The results suggest that the rejection frequencies decrease as the ARCH parameter β becomes larger. However, the test still shows ability to reject the linearity null hypothesis, at least when there is 400 observations at hand.

5 Conclusions

The asymptotic test for smooth transition autoregressive nonlinearities exhibit a seriously distorted size, whereas, the concept of bootstrap testing works extraordinarily well. If the significance level is calculated by a bootstrap procedure an exact test is almost always the result. However, the resampling algorithm must incorporate conditional heteroskedasticity when the true disturbance process in governed by ARCH. When the errors are *iid* normally distributed the choice of resampling scheme is not very important.

The simulation results suggest that the bootstrap testing procedure has power against STAR alternatives, even if the disturbance process is conditionally heteroskedastic.

If prior information, such as theory or test results, suggests that the series at hand do not have ARCH, the asymptotic test has nice size properities and highest possible power. Otherwise, the test requires a ARCH preserving bootstrap.

Acknowledgements

Financial support to Michael Andersson from the Swedish Research Council for Humanities and Social Sciences (HSFR) and to Johan Lyhagen from the Tore Browaldhs Foundation for Scientific Research and Education is gratefully acknowledged.

References

Davidson, R. and J.G. MacKinnon, 1996, The power of bootstrap tests. Queens Institute for Economic Research Discussion Paper No. 937.

Davidson, R. and J.G. MacKinnon, 1999a, The size distortion of bootstrap tests. Forthcomming Econometric Theory.

Davidson, R. and J.G. MacKinnon, 1999b, Bootstrap tests: How many bootstraps? Forthcomming Econometric Reviews.

Efron, B. and R.J. Tibshirani, 1993, An introduction to the bootstrap. Chapman & Hall, New York.

Lundbergh, S. and T. Teräsvirta, 1998, Modelling economic high-frequency time series with STAR-STGARCH models, Working Papers in Economics and Finance, No 291, Stockholm School of Economics.

Teräsvirta, T. 1994a, Specification, estimation, and evaluation of smooth transition autoregressive models, Journal of the American Statistical Association 89, 208-218.

Teräsvirta, T. 1994b, Testing linearity and modelling nonlinear time series, Kybernetika 30, 319-330.

Table 1: Rejection percentage of the STAR tests when the data follow an AR(1) process with normal errors.

	ϕ						
Test	-0.9	-0.6	-0.2	0.0	0.2	0.6	0.9
$T = 50$							
Asy	**5.41**	4.87	4.79	**4.29**	**4.21**	**3.79**	5.03
B_S	4.76	4.76	5.14	5.01	4.90	4.66	5.34
B_A	**4.00**	**3.45**	**3.36**	**3.23**	**3.07**	**3.12**	**3.91**
$T = 100$							
Asy	5.39	5.08	4.74	**4.31**	**4.47**	**3.90**	**4.54**
B_S	4.81	4.79	4.91	4.65	5.04	4.89	4.74
B_A	**4.17**	**3.45**	**3.21**	**3.88**	**3.86**	**3.46**	**3.81**
$T = 200$							
Asy	5.04	**5.70**	5.21	4.85	5.18	**4.42**	4.61
B_S	5.01	4.61	5.01	4.94	5.05	5.20	5.15
B_A	**4.47**	**4.19**	**3.28**	**3.43**	**3.63**	**4.10**	**4.43**

The number reported is the rejection percentage of the linearity test. Bold face denotes a significant deviation from the nominal (5%) size, based on the 95% acceptance interval $(4.6, 5.4)$. *Asy* denotes the asymptotic test and B_S and B_A the bootstrap testing procedures respectively.

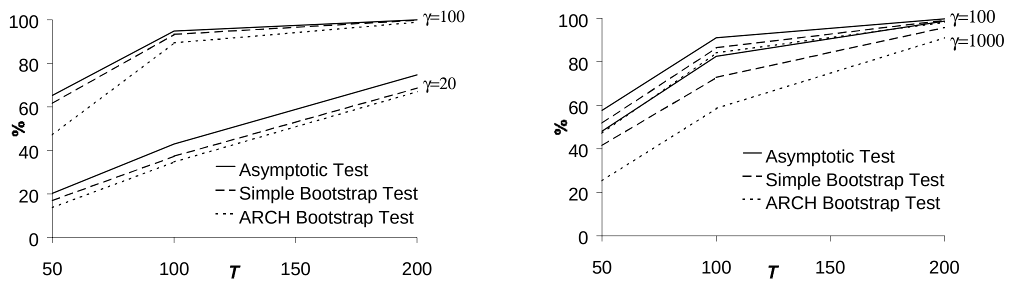
Table 2: Rejection percentage of the STAR tests when the data follow an AR(1) process with ARCH(1) errors.

	ϕ						
Test	-0.9	-0.6	-0.2	0.0	0.2	0.6	0.9
$T = 50$				$\beta = 0.5$			
Asy	**17.30**	**22.12**	**25.24**	**25.04**	**25.03**	**21.06**	**18.24**
B_S	**13.69**	**21.25**	**26.13**	**26.05**	**25.94**	**22.19**	**17.45**
B_A	5.08	5.28	4.91	5.16	5.19	5.07	4.81
$T = 100$							
Asy	**22.26**	**31.59**	**37.09**	**35.99**	**36.00**	**29.99**	**23.73**
B_S	**19.79**	**30.84**	**36.83**	**36.56**	**37.06**	**31.46**	**23.00**
B_A	**4.47**	4.61	4.78	**4.52**	**4.54**	**4.36**	**4.26**
$T = 200$							
Asy	**29.47**	**42.23**	**49.91**	**48.17**	**48.12**	**40.49**	**30.48**
B_S	**27.63**	**41.50**	**49.42**	**48.15**	**48.22**	**42.20**	**29.8**
B_A	4.62	4.74	4.56	4.85	4.71	**4.47**	**4.43**
$T = 50$				$\beta = 0.9$			
Asy	**32.35**	**39.77**	**45.05**	**45.47**	**44.72**	**41.00**	**38.00**
B_S	**25.76**	**37.96**	**44.61**	**45.64**	**45.25**	**41.06**	**32.72**
B_A	**6.02**	**6.19**	**6.96**	**6.77**	**7.08**	**6.31**	**6.22**
$T = 100$							
Asy	**47.83**	**57.86**	**63.44**	**62.23**	**61.69**	**56.72**	**51.45**
B_S	**40.23**	**55.47**	**62.48**	**62.19**	**62.02**	**57.14**	**46.63**
B_A	5.04	**5.58**	**5.47**	**5.59**	**5.68**	**5.44**	4.95
$T = 200$							
Asy	**64.31**	**73.49**	**77.66**	**77.49**	**76.84**	**73.67**	**65.79**
B_S	**58.04**	**72.05**	**76.68**	**76.91**	**76.55**	**73.31**	**60.34**
B_A	5.33	5.09	4.69	5.02	5.03	4.73	4.99

See note to Table 1. The errors follow an ARCH process with parameter β.

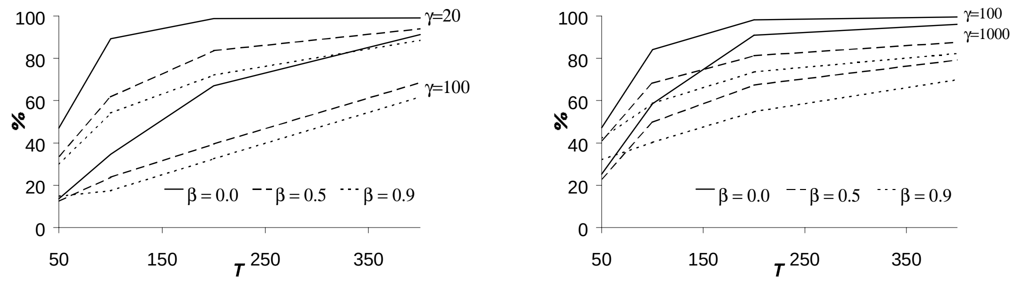
Figure 1: Rejection percentage of the STAR tests when the data follow an ESTAR or an LSTAR process.

LSTAR	*ESTAR*



The figures report the rejection frequencies for the asymptotic and the bootstrap tests. The generated disturbances are normally distributed.

LSTAR	*ESTAR*



The figures report the rejection frequencies for the ARCH bootstrap test B_A. The generated disturbances follow an ARCH process with parameter β.

See note to Figure 1. The data are generated according to equation (8).