

# Department of Economics

## Testing for ARCH in the Presence of Nonlinearity of Unknown Form in the Conditional Mean

Andrew P. Blake and George Kapetanios

Working Paper No. 496

July 2003

ISSN 1473-0278



Queen Mary  
University of London

# Testing for ARCH in the presence of nonlinearity of unknown form in the conditional mean

Andrew P. Blake\*  
CCBS, Bank of England

and

George Kapetanios†  
Queen Mary, University of London

July 1, 2003

## Abstract

Tests of ARCH are a routine diagnostic in empirical econometric and financial analysis. However, it is well known that misspecification of the conditional mean may lead to spurious rejections of the null hypothesis of no ARCH. Nonlinearity is a prime example of this phenomenon. There is little work on the extent of the effect of neglected nonlinearity on the properties of ARCH tests. This paper provides some such evidence and also new ARCH testing procedures that are robust to the presence of neglected nonlinearity. Monte Carlo evidence shows that the problem is serious and that the new methods alleviate this problem to a very large extent.

*Keywords:* Nonlinearity; ARCH; Neural Networks

*JEL Classification:* C12; C22; C45

---

\*Centre for Central Banking Studies, Bank of England, Threadneedle Street, London EC2R 8AH, UK. Email: Andrew.Blake@bankofengland.co.uk.

†Department of Economics, Queen Mary, University of London, Mile End Road, London E1 4NS, UK. Email: G.Kapetanios@qmul.ac.uk

# 1 Introduction

Since the introduction of the autoregressive conditional heteroscedasticity (ARCH) model by Engle (1982) testing for the presence of ARCH has become a routine diagnostic in the econometric analysis of macroeconomic and especially financial time series. A large variety of different tests have been developed. The most usual is based on an autoregression of the squared residual on a constant and its  $p$  lags whereby the joint significance of all the lags included is tested.

As observed by a number of authors, tests for ARCH may reject the null hypothesis of no ARCH if other misspecifications of the conditional mean of the model are present. Notable cases include work by Bera, Higgins, and Lee (1992), Bera and Higgins (1997) and Lumsdaine and Ng (1999). Lumsdaine and Ng (1999) suggest procedures based on recursive residuals that may alleviate misspecifications in the conditional mean and thereby reduce the potential for falsely rejecting the null of no ARCH when other forms of misspecification are present in the model. Bera and Higgins (1997) observe that bilinear and ARCH models have a similar unconditional moment structure raising the possibility that bilinear and ARCH processes may be confused in practical applications.

The current paper follows on from this literature. In particular we suggest that other forms of nonlinearity in the conditional mean may be causing false rejection in testing for ARCH. Examples include nonlinearities of the smooth transition autoregressive (STAR) form or the self-exciting threshold autoregressive (SETAR) form. This observation in itself is not novel. However, there is little work that investigates the interplay of the degree of nonlinearity with the degree of rejection of the no-ARCH null hypothesis. We carry out such investigation using Monte Carlo experiments.

Using this observation as a starting point we suggest new testing procedures that enable valid detection of ARCH even in the presence of nonlinearity. The approach we take is based on neural networks. Neural networks are

a flexible form of nonlinear model that are able to approximate nonlinearities of unknown form (up to the choice of variables that enter the nonlinear conditional mean function) arbitrarily well and thereby produce well behaved residuals under the null of no ARCH on which standard ARCH tests may then be carried out.

However, obtaining the residuals of what are in effect neural network models is not straightforward. Estimation using a standard neural network specification based on the logistic function involves nonlinear least squares. There are additional identification problems if the true DGP happens to be linear. We avoid the need for complicated estimation techniques by using two alternative strategies that also resolve potential identification issues. The first uses an alternative to the logistic neural network model, the radial basis function (RBF) artificial neural network model. This type of neural network has been used in the econometric literature to test for neglected nonlinearity (see Blake and Kapetanios (2000a)). Estimation of such models can be carried out using ordinary least squares. The second strategy uses polynomial approximations of the logistic neural network specification following the work of Teräsvirta, Lin, and Granger (1993). Again, this leads to a model that can be estimated using ordinary least squares.

The plan of the paper is as follows. Section 2 gives more details about the nature of the problem we consider by drawing on previous work in the area. Section 3 discusses the new testing procedures we suggest. Section 4 provides a Monte Carlo investigation of the new and existing testing procedures that reveals the extent of the problem and the ability of the new tests to deal with it. Although this is at a cost of reduced power for the detection of ARCH when it actually is present, it is offset by the enormous gains in producing properly sized tests. Section 5 concludes.

## 2 Testing for ARCH under conditional mean nonlinearity

We concentrate on the following univariate model for the series  $y_t$ ,  $t = 1, \dots, T$ :

$$y_t = f(x_{1,t}, \dots, x_{p,t}; \alpha) + \epsilon_t \quad (1)$$

where  $f(\cdot; \cdot)$  is an unknown function and  $\epsilon_t$  has mean zero, variance conditional on a Borel measurable  $\sigma$ -field with respect to the exogenous (or predetermined at time  $t$ ) variables  $x_{1,t}, x_{2,t}, \dots, x_{p,t}$  denoted by  $\mathcal{I}_{t-1}$ , given by:

$$h_t = \beta_0 + \sum_{j=1}^q \beta_j \epsilon_{t-j}^2 \quad (2)$$

and unconditional variance given by  $\sigma^2$ . The variables  $x_{j,t}$  are assumed to be stationary and ergodic with finite second moments. This setup encompasses both linear and nonlinear autoregressive stationary models. The null hypothesis of interest is:

$$H_0 : \beta_1 = \dots = \beta_q = 0. \quad (3)$$

If one assumes linearity then the conditional mean model becomes:

$$y_t = \alpha_0 + \sum_{j=1}^p \alpha_j x_{j,t} + \epsilon_t. \quad (4)$$

The estimated OLS residual is given by:

$$\hat{\epsilon}_t = \epsilon_t + f(x_{1,t}, \dots, x_{p,t}; \alpha) - \hat{\alpha}_0 - \sum_{j=1}^p \hat{\alpha}_j x_{j,t} = \chi_t + \epsilon_t \quad (5)$$

with the squared residual given by:

$$\begin{aligned} \epsilon_t^2 + (f(x_{1,t}, \dots, x_{p,t}; \alpha) - \hat{\alpha}_0 - \sum_{j=1}^p \hat{\alpha}_j x_{j,t})^2 \\ + 2\epsilon_t(f(x_{1,t}, \dots, x_{p,t}; \alpha) - \hat{\alpha}_0 - \sum_{j=1}^p \hat{\alpha}_j x_{j,t}) = \epsilon_t^2 + \psi_t. \end{aligned} \quad (6)$$

The expectation of the above conditional on  $\mathcal{I}_{t-1}$  is clearly not constant even if the null hypothesis holds. Of course, this analysis holds for any misspecified conditional mean function and the root of the false rejection in ARCH tests, under the null hypothesis, lies in the presence of serial correlation in  $\psi_t$  as Lumsdaine and Ng (1999) have observed. Our conjecture is that nonlinearity in  $\chi_t$  may induce large probabilities of rejection of the null hypothesis of ARCH tests, under the null hypothesis, compared with linear misspecification arising by, say, the omission of an extra lag for the conditional mean model. Clearly, this probability will depend on many things chief among which is the actual values of the parameters of the model and the nonlinearity considered. Therefore, any rigorous theoretical demonstration is bound to be of limited use. A heuristic argument for our conjecture may be given as follows: Assume two cases of conditional mean misspecification. In the first case the conditional mean is assumed to be constant when the true conditional mean model is an AR(1) model. In the second case the assumed model is an AR(1) model but the true model is a SETAR model of the form:

$$y_t = \gamma_0 + \gamma_1 I(y_{t-1} < r)y_{t-1} + \gamma_2 I(y_{t-1} \geq r)y_{t-1} + \epsilon_t \quad (7)$$

where  $I(\cdot)$  is the indicator function. In the first case:

$$\chi_t = (\hat{\alpha}_0 - \alpha_0) + \alpha_1 y_{t-1}. \quad (8)$$

In the second case:

$$\begin{aligned} \chi_t &= (\hat{\gamma}_0 - \gamma_0) + \tilde{\gamma}_1 I(y_{t-1} < r)y_{t-1} + \tilde{\gamma}_2 I(y_{t-1} \geq r)y_{t-1} \\ &= (\hat{\gamma}_0 - \gamma_0)\gamma_{t-1}y_{t-1} \end{aligned} \quad (9)$$

where  $\tilde{\gamma}_i = \hat{\alpha}_1 - \gamma_i$ ,  $i = 1, 2$  and:

$$\gamma_{t-1} = \tilde{\gamma}_1 I(y_{t-1} < r) + \tilde{\gamma}_2 I(y_{t-1} \geq r). \quad (10)$$

In the first case variation in the conditional variance of the residual, under the null of no ARCH, comes from  $y_{t-1}$  only. In the second case both  $\gamma_{t-1}$  and

$y_{t-1}$  contribute to the variation. This example is intended to illustrate the possibility that neglected nonlinearity may induce complex forms of variation in the conditional variance and therefore lead to acute problems of overrejection for standard ARCH tests. We indeed find this to be the case with a variety of nonlinear models in Section 4. It is then reasonable to suggest that methods for accounting for general forms of nonlinearity prior to applying ARCH tests may be useful. The next section suggests such methods.

### 3 Nonlinearity robust ARCH tests

Following on from the previous section it is clear that nonlinearity has the potential to introduce problems in the detection of ARCH for dynamic models. We therefore require a test for ARCH that is robust to a large class of nonlinearities. Artificial neural network models provide an ideal framework for this analysis. This is due to the useful property that they can approximate continuous functions arbitrarily well. More specifically, a continuous function  $f(z)$  can be arbitrarily well approximated in the supremum norm by  $\sum_{i=1}^q b_i g(z'_i)$  for finite  $q$  and  $z'_i = d_{0,i} + d'_i z$  if either (i)  $g(\cdot)$  is sigmoidal, i.e.  $g(\cdot)$  is non-decreasing with  $\lim_{z \rightarrow -\infty} g(z) = 0$  and  $\lim_{z \rightarrow \infty} g(z) = 1$  or (ii)  $g(\cdot)$  has non-zero Lebesgue measure expectation and is  $L_p$  bounded for some  $p \geq 1$ . For more details see Hornik, Stinchcombe, and White (1989), Stinchcombe and White (1989) and Cybenko (1989).<sup>1</sup> The universal approximator properties of neural networks have been put to effect in the econometrics literature by, among others, Lee, White, and Granger (1993) and Blake and Kapetanios (2000a) to construct tests for neglected nonlinearity in stationary models.<sup>2</sup>

In the present context we wish to robustify ARCH tests by fitting a neural

---

<sup>1</sup>See Campbell, Lo, and MacKinlay (1997) for an excellent introduction to artificial neural networks, which covers the RBF networks that we use below. Bishop (1995) gives a thorough and very readable account.

<sup>2</sup>See also Peguin-Feissolle (1999) and Blake and Kapetanios (2000b) for general ARCH testing and Blake and Kapetanios (2003) for an application to unit root testing.

network model to (1), obtaining the residuals and carrying out a standard ARCH test. By the universal approximation property of the neural networks for continuous functions discussed above we know that the model (1) may be written as:

$$y_t = a_0 + a_1'x_t + \sum_{i=1}^q b_i g(d_{0,i} + d_i'x_t) + \epsilon_t \quad (11)$$

where  $x_t = (x_{1,t}, \dots, x_{p,t})'$  with  $g(\cdot)$  now a known function. This specification can then be estimated consistently and the residuals then tested for ARCH using standard tests since  $\hat{\epsilon}_t$  from the estimation of (11) will converge in probability to  $\epsilon_t$ . However estimation of (11), although feasible, may not be easy as it requires nonlinear least squares for the function used for  $g(\cdot)$ . For testing, however, this may not be necessary. There are neural network specifications or approximations of neural networks where such iterative schemes are not needed for estimation. We consider two.

The first is a radial basis function (RBF) artificial neural network, which is often referred to as a linear network. These are simple to define. Let there be  $q$  centers, denoted  $c_j$ , and a radius vector  $\tau$ . Utilise a function that is monotonically decreasing about  $c_j$ . A natural choice is the Gaussian function, and the appropriate RBF is:

$$g([c_j' \ \tau']', x_t) = e^{-\|(x_t - c_j)/\tau\|^2} \quad (12)$$

where the division notation denotes element by element division. By the monotonicity property, each RBF has maximum activation (of unity) when the input vector coincides with the  $j$ th center independent of  $\tau$ . Conversely, if the input vector is far enough away from the center (controlled by  $\tau$ ) the activation is zero.<sup>3</sup> The linearity of this network derives from the property that if the centers ( $c_j$ ) and radii ( $\tau$ ) are determined by some procedure then

---

<sup>3</sup>Other functional forms (such as the multiquadratic) have the same property and can be used instead, although in experiments Blake and Kapetanios (2000a) found very little difference in performance between various functions for the related problem of nonlinearity testing.



the RBFs can be straightforwardly used as additional regressors. The trick is to use data-based procedures to determine both. It is convenient to use a simple function of the data, such as a multiple of the maximum change from period  $t$  to period  $t + 1$  for all  $t$  of each input as the radius for that input. In the previous related work cited above, we used the alternative of a unit variance for the normalised data and found that it worked well. This defines  $\tau$ . Following Orr (1995), we then allow there to be  $T$  potential centers (the  $c_j$ ) for the RBFs which is each of the vector of observations. The RBFs thus obtained are ranked according to their ability to reduce the unexplained variance of (11) when entered individually. Then we successively add the sorted RBFs into (11) until we minimize an information criterion, chosen to be AIC. This procedure is known as forward selection. The pair of data-based procedures (normalisation to set the radii and center choice from the data points by forward selection) yield a linear estimation problem for  $b$  as all the terms in (12) are defined by the process.

A second approach that is amenable to linear estimation is motivated by a test of neglected nonlinearity that approximates the logistic neural network developed by Teräsvirta, Lin, and Granger (1993). The logistic network uses:

$$g(d, x_t) = 1/(1 + e^{-d'x_t}) \quad (13)$$

for the activation function. As remarked above, to estimate the complete nonlinear model (all the parameters  $a$ ,  $b$  and  $d$ ) is costly. Teräsvirta, Lin, and Granger (1993) suggest a third-order Taylor expansion of the logistic neural network is used instead. This choice of expansion is arbitrary, and can clearly be replaced by the  $n$ -th order Taylor expansion of the logistic neural network. This turns out to be an  $n$ -th order polynomial in  $(x_{1,t} + \dots + x_{p,t})$ , where the operational order of the Taylor expansion needs to be chosen. In common with our treatment of the RBF network we use an information criterion, again AIC. This is chosen from expansions of order two, three and four where for simplicity for the third and fourth order Taylor expansions only cross products of (powers of) up to two variables are considered.

Note that there is a problem here that the model we estimate is not necessarily appropriate for a linear model. Indeed, for a linear model the additional terms should be absent. This identification problem was solve by Lee, White, and Granger (1993) by adding a fixed number randomly generated logistic functions as regressors. Our two procedures solve it by the use of the information criterion, which is restricted to pick a minimum of one additional regressor. Our test for ARCH, of course, uses the residuals from this model.

After fitting the (approximate) neural network model, the residuals obtained can be tested using any residual-based ARCH test. The asymptotic distribution of the usual LM ARCH test is still  $\chi_q^2$  under the null hypothesis and assuming no nonlinearity. Further refinement of the new tests may be contemplated if the techniques suggested by Lumsdaine and Ng (1999) are combined with the neural network specifications. In particular if, as suggested by Lumsdaine and Ng (1999), (functions of) lags of recursive residuals from the original linear specification in (1), help in picking up misspecifications in the conditional mean then augmenting the variable set  $x_t$  by (functions of) this lagged recursive residual may further improve the performance of the tests under the null hypothesis. Additionally, OLS residuals may be used in place of recursive residuals. In small samples these will introduce biases since they depend on the whole of the sample via the parameter estimates but asymptotically these parameter estimates will converge to some limit and the lagged OLS residuals will not cause any further asymptotic biases.

## 4 Monte Carlo experiments

In this section we carry out Monte Carlo experiments to illustrate the problem arising out of pronounced nonlinearity in the context of ARCH testing and the extent to which the new tests alleviate the problem. The issue occurs

under the null hypothesis and therefore the size performance of the existing tests is under scrutiny. Nevertheless, it is important that the new methods do not reduce the power of the ARCH tests unduly. Therefore, there is the usual tradeoff between power and size which needs to be explored.

We concentrate on nonlinear autoregressive models and add a linear AR(1) model given by  $y_t = \alpha y_{t-1} + \epsilon_t$  for comparison. We consider three different classes of nonlinear models under the null hypothesis. These are SETAR, STAR and bilinear models. The models are given by:

$$y_t = \gamma_1 I(y_{t-1} < r) y_{t-1} + \gamma_2 I(y_{t-1} \geq r) y_{t-1} + \epsilon_t \quad (14)$$

$$y_t = \delta_1 y_{t-1} + \delta_2 (1 - e^{-\delta_3 y_{t-1}^2}) y_{t-1} + \epsilon_t \quad (15)$$

and:

$$y_t = \zeta \epsilon_{t-1} y_{t-1} + \epsilon_t \quad (16)$$

Under the null hypothesis of no ARCH both the conditional and unconditional variance of  $\epsilon_t$  is equal to  $\sigma^2$  which is set to 1 for all size experiments. We consider one DGP from the AR class, four DGPs from SETAR models, four DGPs from STAR models and three DGPs from bilinear models. The coefficients for each class are:

- AR Model
  - Experiment 1:  $\alpha = 0.5$ .
- SETAR Models
  - Experiment 2:  $\gamma_1 = 0.1, \gamma_2 = 0.5, r = 0$ .
  - Experiment 3:  $\gamma_1 = -0.1, \gamma_2 = 0.5, r = 0$ .
  - Experiment 4:  $\gamma_1 = -0.3, \gamma_2 = 0.9, r = 0$ .
  - Experiment 5:  $\gamma_1 = -0.8, \gamma_2 = 0.9, r = 0$ .
- STAR Models

- Experiment 6:  $\delta_1 = 0.5, \delta_2 = -0.5, \delta_3 = 0.05$ .
- Experiment 7:  $\delta_1 = 0.5, \delta_2 = -0.5, \delta_3 = 0.5$ .
- Experiment 8:  $\delta_1 = 0.9, \delta_2 = -1.4, \delta_3 = 0.05$ .
- Experiment 9:  $\delta_1 = 0.9, \delta_2 = -1.4, \delta_3 = 0.5$ .

- Bilinear Models

- Experiment 10:  $\zeta = 0.1$ .
- Experiment 11:  $\zeta = 0.3$ .
- Experiment 12:  $\zeta = 0.7$ .

For experiments relating to power we consider the following ARCH model. The AR(1) model above is used for the conditional mean and an ARCH(1) model of  $h_t = \beta_0 + \beta_1 \epsilon_t^2$  for the conditional variance. For each of three experiments the parameters are:

- ARCH Models

- Experiment 13:  $\alpha = 0.5, \beta_0 = 0.1, \beta_1 = 0.1$ .
- Experiment 14:  $\alpha = 0.5, \beta_0 = 0.1, \beta_1 = 0.5$ .
- Experiment 15:  $\alpha = 0.5, \beta_0 = 0.1, \beta_1 = 0.9$ .

All errors are obtained from the GAUSS normal pseudo-random number generator. For every sample, initial conditions are set to zero and 20 observations are dropped to minimise dependence on the choice of initial conditions. Throughout the lag order,  $p$ , is set to 1. We consider samples of 100 and 200 observations.

We consider two ARCH tests: The first is the usual LM-test (and denoted LM in the tables) and the second is the ARCH test developed by Peguin-Feissolle (1999) which is a residual-based test using neural network ideas derived from the work of Lee, White, and Granger (1993) (denoted NN in

the tables). We choose this test because we wish to illustrate that radically different ARCH tests suffer from the problem of overrejection under the null to a similar extent and can benefit from the approach we suggest. We have two possible nonlinear modelling choices. The subscript *tlg* denotes the use of the Taylor expansion method and the subscript *rbf* denotes use of the RBF neural network. Also we consider two extended ARCH tests. The superscript *r* denotes augmentation of the variable set by the lag of the recursive residual and its square, as suggested by Lumsdaine and Ng (1999), whereas the superscript *o* denotes similar augmentation of the variable set only using the lag of the OLS residual. Results in the form of rejection probabilities are presented in Tables 1–4. We do not report size corrected powers because it is not clear what are the proper empirical critical values to use to correct the rejection probabilities under the alternative. As we will see the rejection probabilities under the null of no ARCH vary greatly depending on the nonlinear model used.

The results make very interesting reading. The performance of the *LM* and *NN* tests depend markedly on the nonlinear model used. Rejection probabilities under the null reach level of 90% for the bilinear model, which was expected given the work of Bera and Higgins (1997), but even for SETAR and STAR models pronounced nonlinearity, measured by the difference between  $\gamma_1$  and  $\gamma_2$  for SETAR models and the magnitude of  $\delta_2$  for STAR models, induces rejection probabilities of up to 50%. Clearly nonlinearity in the mean and ARCH are difficult to distinguish using standard tests. The recursive residual tests of Lumsdaine and Ng (1999) are considerable improvements on the standard tests. Rejection probabilities fall substantially but still remain quite high for very nonlinear conditional mean processes. For example for the most extreme SETAR model, experiment 5, and 200 observations the rejection probability for the LM test is 26%.

Moving on to the new testing procedures we observe a dramatic improvement. Rejection probabilities are very close to 5%. None of the SETAR or

STAR nonlinear models can induce even minor departures from the correct significance level. Bilinear models still produce overrejections. However this is to be expected: The variables involved in the bilinear model include the lagged error. This is not included in the set of variables used in the construction of the neural network. Therefore, the nonlinearity cannot be completely captured. However, even for these models noticeable and worthwhile improvement is observed. In particular adding recursive or OLS residuals produces satisfactory rejection probabilities apart from the final Bilinear model, experiment 11, where the nonlinearity is very pronounced.

When we consider the rejection probabilities under the alternative we see that the new procedures do not exhibit a very marked reduction in these probabilities, as desired. In particular we observe a reduction of about 20% for tests that do not include recursive or OLS residuals for samples of 100 observations and only a reduction of about 10% for samples of 200 observations. In some cases, as in Table 4, there is almost no loss of power for the  $NN_{tlg}$  and  $NN_{rbf}$  tests. The results are indeed very encouraging and clearly suggest that the new methods are useful.

## 5 Conclusions

It is well known that tests for ARCH are powerful against a wider variety of misspecifications. In particular it is well known from the work of Lumsdaine and Ng (1999) and others that misspecification in the conditional mean may lead to spurious rejection of the no ARCH hypothesis. However, apart from the general heuristic methods of Lumsdaine and Ng (1999) there is little in terms of methods to avoid this problem.

This paper tries to suggest a solution in the case where the conditional mean function suffers from neglected nonlinearity of unknown form. We suggest the use of neural networks to approximate to an arbitrary level of accuracy the unknown nonlinearity and having removed it, test the residuals

for ARCH using standard tests. Monte Carlo evidence has suggested that the new methods are able to remove the large distortions introduced by nonlinearity at a rather modest cost in terms of power loss.

## References

- BERA, A. K., AND M. L. HIGGINS (1997): “ARCH and Bilinearity as Competing Models for Conditional Dependence,” *Journal of Business and Economic Statistics*, 15(1), 43–50.
- BERA, A. K., M. L. HIGGINS, AND S. LEE (1992): “Interaction Between Autocorrelation and Conditional Heteroscedasticity: A Random-Coefficient Approach,” *Journal of Business and Economic Statistics*, 10(2), 133–142.
- BISHOP, C. M. (1995): *Neural Networks for Pattern Recognition*. Oxford University Press.
- BLAKE, A. P., AND G. KAPETANIOS (2000b): “A Radial Basis Function Artificial Neural Network Test for ARCH,” *Economics Letters*, 69(1), 15–23.
- (2000a): “A Radial Basis Function Artificial Neural Network Test for Neglected Nonlinearity,” Working Paper 153, National Institute of Economic and Social Research.
- (2003): “Pure Significance Tests of the Unit Root Hypothesis Against Nonlinear Alternatives,” *Journal of Time Series Analysis*, 24(3), 253–267.
- CAMPBELL, J. Y., A. W. LO, AND A. C. MACKINLAY (1997): *The Econometrics of Financial Markets*. Princeton University Press.
- CYBENKO, G. (1989): “Approximation by Superpositions of a Sigmoidal Function,” *Mathematics of Control, Signals and Systems*, 2, 304–314.
- ENGLE, R. F. (1982): “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation,” *Econometrica*, 50(1), 987–1007.

- HORNIK, K., M. STINCHCOMBE, AND H. WHITE (1989): “Multilayer Feed-forward Networks are Universal Approximators,” *Neural Networks*, 2, 359–366.
- LEE, T. H., H. WHITE, AND C. W. J. GRANGER (1993): “Testing for Neglected Nonlinearity in Time Series Models: A Comparison of Neural Network Methods and Alternative Tests,” *Journal of Econometrics*, 56(3), 269–290.
- LUMSDAINE, R. L., AND S. NG (1999): “Testing for ARCH in the Presence of a Possibly Misspecified Conditional Mean,” *Journal of Econometrics*, 93(2), 257–279.
- ORR, M. J. (1995): “Regularisation in the Selection of Radial Basis Function Centers,” *Neural Computation*, 7(3), 606–623.
- PEGUIN-FEISSOLLE, A. (1999): “A Comparison of the Power of Some Tests for Conditional Heteroscedasticity,” *Economics Letters*, 63(1), 5–17.
- STINCHCOMBE, M., AND H. WHITE (1989): “Universal Approximation Using Feedforward Networks with Non-Sigmoid Hidden Layer Activation Functions,” *Proceedings of the International Joint Conference on Neural Networks*, 1, 359–366.
- TERÄSVIRTA, T., C. F. LIN, AND C. W. J. GRANGER (1993): “Power of the Neural Network Linearity Test,” *Journal of Time Series Analysis*, 14, 209–220.



Table 1: Monte Carlo Results for  $T = 100$  and ARCH LM tests

		$LM$	$LM^r$	$LM_{tlg}$	$LM_{rbf}$	$LM_{tlg}^r$	$LM_{rbf}^r$	$LM_{tlg}^o$	$LM_{rbf}^o$
					Size				
AR	1	0.050	0.037	0.029	0.027	0.031	0.034	0.027	0.024
SETAR	2	0.043	0.026	0.027	0.026	0.023	0.027	0.029	0.025
	3	0.049	0.039	0.031	0.035	0.024	0.027	0.031	0.029
	4	0.144	0.081	0.029	0.028	0.023	0.031	0.031	0.027
	5	0.280	0.130	0.023	0.026	0.019	0.023	0.032	0.027
STAR	6	0.048	0.040	0.037	0.036	0.027	0.031	0.026	0.031
	7	0.050	0.037	0.044	0.041	0.042	0.042	0.031	0.041
	8	0.137	0.088	0.042	0.046	0.026	0.034	0.036	0.027
	9	0.097	0.077	0.031	0.033	0.022	0.034	0.024	0.026
Bilin	10	0.051	0.030	0.026	0.023	0.021	0.026	0.023	0.021
	11	0.458	0.127	0.046	0.046	0.037	0.033	0.018	0.032
	12	0.905	0.773	0.471	0.410	0.194	0.299	0.099	0.307
					Power				
ARCH	13	0.112	0.085	0.074	0.074	0.036	0.044	0.039	0.057
	14	0.701	0.567	0.491	0.464	0.246	0.314	0.236	0.333
	15	0.887	0.764	0.643	0.627	0.407	0.488	0.324	0.524

Table 2: Monte Carlo Results for  $T = 200$  and ARCH LM tests

		$LM$	$LM^r$	$LM_{tlg}$	$LM_{rbf}$	$LM_{tlg}^r$	$LM_{rbf}^r$	$LM_{tlg}^o$	$LM_{rbf}^o$
		Size							
AR	1	0.047	0.037	0.037	0.040	0.027	0.032	0.033	0.033
SETAR	2	0.055	0.047	0.044	0.041	0.042	0.045	0.046	0.043
	3	0.064	0.039	0.042	0.040	0.033	0.041	0.034	0.031
	4	0.258	0.124	0.036	0.037	0.035	0.038	0.047	0.041
	5	0.487	0.266	0.044	0.045	0.042	0.039	0.042	0.039
STAR	6	0.044	0.039	0.036	0.036	0.039	0.047	0.034	0.035
	7	0.050	0.042	0.025	0.027	0.026	0.029	0.033	0.032
	8	0.208	0.150	0.041	0.041	0.041	0.041	0.046	0.042
	9	0.221	0.184	0.037	0.039	0.028	0.033	0.035	0.030
Bilin	10	0.080	0.030	0.030	0.027	0.024	0.025	0.028	0.027
	11	0.777	0.173	0.071	0.066	0.043	0.051	0.034	0.047
	12	0.998	0.972	0.834	0.783	0.580	0.716	0.372	0.666
		Power							
ARCH	13	0.255	0.204	0.175	0.176	0.114	0.132	0.121	0.141
	14	0.955	0.915	0.887	0.874	0.749	0.799	0.715	0.814
	15	0.988	0.977	0.917	0.916	0.854	0.882	0.811	0.890

Table 3: Monte Carlo Results for  $T = 100$  and ARCH NN tests

		$NN$	$NN^r$	$NN_{tlg}$	$NN_{rbf}$	$NN_{tlg}^r$	$NN_{rbf}^r$	$NN_{tlg}^o$	$NN_{rbf}^o$
		Size							
AR	1	0.052	0.046	0.039	0.036	0.025	0.029	0.026	0.029
SETAR	2	0.057	0.033	0.023	0.024	0.020	0.020	0.021	0.019
	3	0.053	0.034	0.020	0.022	0.019	0.020	0.013	0.021
	4	0.163	0.085	0.027	0.029	0.016	0.021	0.022	0.023
	5	0.301	0.146	0.026	0.029	0.027	0.025	0.026	0.018
STAR	6	0.063	0.036	0.035	0.042	0.023	0.025	0.023	0.031
	7	0.058	0.036	0.024	0.021	0.032	0.030	0.024	0.031
	8	0.156	0.088	0.023	0.025	0.024	0.022	0.026	0.027
	9	0.136	0.106	0.028	0.029	0.015	0.022	0.012	0.016
Bilin	10	0.072	0.034	0.025	0.025	0.025	0.022	0.018	0.028
	11	0.451	0.128	0.054	0.051	0.023	0.040	0.021	0.036
	12	0.957	0.869	0.643	0.585	0.272	0.436	0.144	0.387
		Power							
ARCH	13	0.130	0.096	0.087	0.089	0.055	0.062	0.050	0.070
	14	0.734	0.654	0.615	0.577	0.378	0.438	0.352	0.444
	15	0.945	0.900	0.838	0.817	0.628	0.720	0.542	0.721

Table 4: Monte Carlo Results for  $T = 200$  and ARCH NN tests

		$NN$	$NN^r$	$NN_{tlg}$	$NN_{rbf}$	$NN_{tlg}^r$	$NN_{rbf}^r$	$NN_{tlg}^o$	$NN_{rbf}^o$
		Size							
AR	1	0.048	0.038	0.027	0.029	0.029	0.024	0.028	0.028
SETAR	2	0.066	0.039	0.035	0.029	0.031	0.026	0.023	0.029
	3	0.077	0.038	0.026	0.028	0.026	0.023	0.021	0.023
	4	0.282	0.128	0.045	0.038	0.025	0.028	0.026	0.025
	5	0.502	0.264	0.036	0.038	0.024	0.030	0.021	0.033
STAR	6	0.048	0.038	0.032	0.032	0.025	0.022	0.028	0.025
	7	0.065	0.047	0.014	0.014	0.027	0.017	0.023	0.016
	8	0.240	0.178	0.031	0.033	0.023	0.023	0.031	0.026
	9	0.275	0.213	0.035	0.038	0.031	0.043	0.023	0.037
Bilin	10	0.105	0.030	0.025	0.024	0.024	0.021	0.021	0.023
	11	0.761	0.181	0.091	0.087	0.052	0.065	0.026	0.063
	12	1.000	0.989	0.929	0.894	0.698	0.852	0.445	0.800
		Power							
ARCH	13	0.238	0.176	0.160	0.152	0.108	0.117	0.115	0.126
	14	0.969	0.949	0.944	0.938	0.850	0.885	0.842	0.893
	15	0.999	0.995	0.985	0.973	0.961	0.975	0.931	0.974

**This working paper has been produced by  
the Department of Economics at  
Queen Mary, University of London**

**Copyright © 2003 Andrew P. Blake and George Kapetanios  
All rights reserved.**

**Department of Economics  
Queen Mary, University of London  
Mile End Road  
London E1 4NS  
Tel: +44 (0)20 7882 5096 or Fax: +44 (0)20 8983 3580  
Email: [j.conner@qmul.ac.uk](mailto:j.conner@qmul.ac.uk)  
Website: [www.econ.qmul.ac.uk/papers/wp.htm](http://www.econ.qmul.ac.uk/papers/wp.htm)**