
Episodic nonstationarity in exchange rates

CHRISTOPHER BROOKS and *MELVIN J. HINICH

ISMA Centre, Department of Economics, The University of Reading, Whiteknights, Reading RG6 6AA, UK and *Applied Research Laboratories and the Department of Government, The University of Texas at Austin, Austin, TX 78713–8029, USA

Received 13 May 1997

We examine a method recently proposed by Hinich and Patterson (mimeo, University of Texas at Austin, 1995) for testing the validity of specifying a GARCH error structure for financial time series data in the context of a set of ten daily Sterling exchange rates. The results demonstrate that there are statistical structures present in the data that cannot be captured by a GARCH model, or any of its variants. This result has important implications for the interpretation of the recent voluminous literature which attempts to model financial asset returns using this family of models.

I. INTRODUCTION

Interest in the class of autoregressive conditionally heteroscedastic models, originally suggested by Engle (1982), to model financial time series has been phenomenal. This class of models relaxes the assumption of the classical linear regression model that the variance of the disturbance term is conditionally as well as unconditionally constant. Since the generalization due to Bollerslev (1986), the number of empirical and theoretical developments in the field has exploded, with rapid development of applications and variants. However, although these models have been found to accurately describe a number of the important characteristics of the data from a diversity of financial disciplines, as Hall *et al.* (1989) note, the ARCH parameterization of the conditional variance does not have any solid grounding in economic theory, but represents 'a convenient and parsimonious representation of the data.'

We can write the standard GARCH-(1,1) model as suggested in Bollerslev (1986):

$$\begin{aligned}\varepsilon_t | \psi_{t-1} &\sim N(0, h_t) \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}\end{aligned}\quad (1)$$

The question of the validity of this formulation in terms of whether it represents an accurate characterization of the data has, however, hardly received a mention in the literature. The purpose of this paper is to concentrate on two aspects of GARCH models: the assumption of strict stationarity, which can be tested using an application of the Hinich third order portmanteau statistic, and the noticeable absence of GARCH

over relatively long sub-periods of the series, using Engle's LM test.

II. THE DATA

The analysis presented here is based on just over 20 years of daily mid-price spot exchange rate data denominated in Sterling. The sample period taken covers the entire post-Bretton Woods era, specifically from 2 January 1974 until 1 July 1994 inclusive. A set of ten currencies are analysed, namely the Austrian schilling, the Canadian dollar, the Danish krone, the French franc, the German mark, the Hong Kong dollar, the Italian lira, the Japanese yen, the Swiss franc, and the US dollar. The raw exchange rates are transformed into a series of 5191 log-returns which can be interpreted as a series of continuously compounded daily returns (Brock *et al.*, 1991). The data are split into a set of 148 nonoverlapping windows of length 35 observations (i.e. approximately seven trading weeks). This window length should be sufficiently long to validly apply the tests and yet short enough for the data generating process to have remained roughly constant. In any case, the results are not sensitive to reasonably large changes in this parameter.

III. THE HINICH PORTMANTEAU BICORRELATION TEST

Hinich (1995) has recently suggested a modified version of the Box and Pierce (1970) portmanteau Q-statistic for autocorrelation and a third order portmanteau statistic, which can in a

sense be viewed as a time domain analogue of the bispectrum test. A full theoretical derivation of the test statistics and a number of Monte Carlo simulations to assess their size and power are given in Hinich (1995), and Hinich and Patterson (1995). We thus state without derivation that the test statistics, denoted C and H respectively, are defined as

$$C = \sum_{r=1}^L \left[C^2(r)/(T-r-1) \right] \sim \chi^2(L) \quad (2)$$

where

$$C(r) = \sum_{k=1}^{T-s} Z(t_k)Z(t_{k+r}),$$

and

$$H = \sum_{s=2}^L \sum_{r=1}^{s-1} \left[G^2(r,s)/(T-s) \right] \sim \chi^2((L-1)L/2) \quad (3)$$

where

$$G(r,s) = \sum_{k=1}^{T-s} \left[Z(t_k)Z(t_{k+r})Z(t_{k+s}) \right], Z(t_k)$$

are the standardized observations (created by subtracting the sample mean of the window, and dividing by its standard deviation) at time $t = k, L = T^c (0 < c < 0.5)$; c is a parameter under the choice of the user. Based on the results of Monte Carlo simulations, the use of $c = 0.4$ is recommended in order to maximize the power of the test whilst ensuring a valid approximation to the asymptotic theory. The test is of a null of pure white noise against an alternative that the process has M non-zero correlations or bicorrelations in the set $0 < r < s \leq L$, i.e. that there exists second or third order dependence in the data generating process, and relies on the property of pure noise that it has zero bicovariance. The test is particularly useful in detecting nonlinear dependencies, since

it has much better small-sample properties, and does not have such stiff data requirements as many of its competitors, such as the BDS test (Brock *et al.*, 1987; see Brock *et al.*, 1991 for a useful survey).

It is possible to use the test described above to test whether a GARCH formulation represents an adequate characterization of the data. This is achieved by transforming the returns into a set of binary data using $\{y(t)\} : y(t) = 1$ if $Z(t) \geq 0$ and $y(t) = -1$ if $Z(t) < 0$. If the original $Z(t)$ are generated by a pure ARCH or GARCH process, then $\{y(t)\}$ will be a stationary pure noise series provided that the innovations have a symmetric density with zero mean.¹ The binary transformed data has moments which are well-behaved with respect to the asymptotic theory. Define a window as significant if either the C or H statistic rejects the null of whiteness at the specified threshold level. If the data were generated by any kind of ARCH or GARCH model alone, only that number of windows equal to the actual size of the test will be significant. We can apply the test either to the raw returns or to the residuals of an autoregressive fit of the data. The result of applying the test to the 148 windows for the residuals of an AR(4) model for each series, using a threshold for significance of 0.1% are shown in the following table. The justification for considering the residuals is to demonstrate that the nonstationarity must be a consequence of nonlinearity which is episodically present in the data rather than a form of linear dependence (which has been removed); thus only significant H statistics will cause a rejection of the null of pure noise.

The results of Table 1 clearly show that a larger number of windows are significant than one would expect purely by chance given the very strict 0.1% nominal threshold level. Therefore the data cannot be generated by a strongly stationary white noise process as required by the GARCH model. This result corroborates that of Brooks (1996) who finds that the GARCH model is unable to explain all of the observed nonlinearity in the foreign exchange data. In order to further verify this result, we simulate the estimated GARCH

Table 1. Number and percentage of significant windows of the binary transformed data

Series	Number of sig. windows	Percentage of sig. windows	Dates of sig. windows
Austrian schilling	1	0.7	27 Jan 76 – 15 Mar 76
Canadian dollar	2	1.4	27 Jan 76 – 15 Mar 76; 21 Aug 84 – 9 Oct 84
Danish krone	2	1.4	27 Jan 76 – 15 Mar 76; 6 Aug 81 – 24 Sep 81
French franc	2	1.4	29 Oct 74 – 16 Dec 74; 2 Aug 77 – 20 Sep 77
German mark	2	1.4	27 Jan 76 – 15 Mar 76; 6 Aug 81 – 24 Sep 81
Hong Kong dollar	2	1.4	27 Jan 76 – 15 Mar 76; 25 June 76 – 12 Aug 76
Italian lira	1	0.7	18 Dec 92 – 27 Jan 93
Japanese yen	3	2.0	27 Jan 76 – 15 Mar 76; 2 Nov 82 – 20 Dec 82; 21 Dec 91 – 31 Jan 92
Swiss franc	2	1.4	29 Oct 74 – 16 Dec 74; 27 Jan 76 – 15 Mar 76
US dollar	2	1.4	27 Jan 76 – 15 Mar 76; 21 Aug 84 – 9 Oct 84

¹ The residuals of a linear autoregressive model will have a mean of approximately zero, and the series used here show a surprising degree of symmetry, although they are leptokurtic.

model for the Swiss franc and run the test using exactly the same conditions as above on the simulated data. The values of the parameters used are $\alpha_0 = 0.0108$; $\alpha_1 = 0.1244$; $\beta_1 = 0.8516$. When we simulate the series 50 000 times, we find that the percentage of significant H statistics is only 0.0036, considerably lower than the nominal 0.1% threshold and the values observed using the actual data.

It is also potentially of interest to consider the dates when these transient epochs of nonlinear dependence occurred. These are given in the last column of Table 1. It is apparent that there is very much more action during the 1970s than in the latter half of the sample period, and in particular, the null of pure noise is rejected for the window comprising 27 January 1976–15 March 1976 for all of the currencies except the Italian lira. The Italian foreign exchanges were closed from 21 January until 1 March 1976 (source: *National Institute Economic Review* no. 76, May 1976).

IV. ARCH EVERYWHERE?

In this section, results on the parameter instability of GARCH models and the transient nature of ARCH effects are given. A test for the presence of ARCH in the residuals of a linear model was suggested by Engle (1982), and can be calculated as follows. First, run a linear regression, saving the residuals, $\hat{\varepsilon}_t$. Then square the residuals, and regress them on p own lags to test for ARCH of order p . The LM test statistic is defined as TR^2 (the number of observations multiplied by the coefficient of multiple correlation) from the last regression, and is distributed as a $\chi^2(p)$ under the null of no order p ARCH effects.

The test has been frequently applied in the literature, but a slightly different approach used here is to employ the test over a set of relatively short nonoverlapping windows of length 200, 400, and 800 (giving 25, 12, and 6 windows respectively), in an attempt to discover whether there is strong evidence of ARCH over all time periods, or whether ARCH is present only for short periods of time. The use of short windows is a useful technique for observing dependencies in time series data which may be masked over longer periods. Table 2 below shows the number of nonsignificant windows using these subsample sizes at the 10%, 5%, and 1% levels of significance.

When the data are split into subsamples, it becomes apparent that there are extremely long periods during which there is no evidence of ARCH effects, but these become absorbed into periods when there is strong evidence of ARCH so that the null of no ARCH is rejected more and more convincingly as the subsamples are aggregated. The most extreme case is that of the US dollar, for which 72% of the subsamples of length 200 show no significant evidence of ARCH at the 5% level, and one third of the samples of length 400 show no evidence of ARCH, even at the 10% level. This result cannot, however,

Table 2. The number of nonsignificant windows using Engle's test for ARCH

Series	Length	Number (percentage) of non-significant windows at		
		10%	5%	1%
Austrian schilling	200	14(56%)	15(60%)	18(72%)
	400	2(17%)	4(33%)	5(42%)
	800	0(0%)	1(17%)	1(17%)
Canadian dollar	200	18(72%)	18(72%)	20(80%)
	400	5(42%)	5(42%)	9(75%)
	800	2(33%)	2(33%)	2(33%)
Danish krone	200	13(52%)	15(60%)	17(68%)
	400	4(33%)	7(58%)	8(67%)
	800	1(17%)	1(17%)	1(17%)
French franc	200	12(48%)	14(56%)	17(68%)
	400	5(42%)	5(42%)	5(42%)
	800	0(0%)	1(17%)	2(33%)
German mark	200	14(56%)	16(64%)	17(68%)
	400	3(25%)	3(25%)	5(42%)
	800	0(0%)	0(0%)	1(17%)
Hong Kong dollar	200	12(48%)	14(56%)	16(64%)
	400	1(8%)	2(17%)	2(17%)
	800	0(0%)	0(0%)	1(17%)
Italian lira	200	16(64%)	17(68%)	18(72%)
	400	3(25%)	3(25%)	7(58%)
	800	0(0%)	0(0%)	1(17%)
Japanese yen	200	13(52%)	14(56%)	18(72%)
	400	4(33%)	4(33%)	7(58%)
	800	1(17%)	1(17%)	1(17%)
Swiss franc	200	12(48%)	14(56%)	19(76%)
	400	2(17%)	2(17%)	2(17%)
	800	0(0%)	0(0%)	0(0%)
US dollar	200	17(68%)	18(72%)	20(80%)
	400	4(33%)	4(33%)	6(50%)
	800	2(33%)	2(33%)	3(50%)

be attributed to a fall in the power of the test at smaller samples, for Engle *et al.* (1985) show that the LM test has reasonable power even for samples smaller than 100 observations, and that the power increases only marginally once the sample size is increased to those investigated here.

V. CONCLUSIONS

It has been shown that the GARCH model cannot be considered a full representation of the process generating financial market returns. In particular, the GARCH model fails to capture the time-varying nature of market returns, and treats coefficients as fixed and being drawn from only one regime. Sterling exchange rates have been shown to be characterized by transient epochs of dependence surrounded by long periods of white noise. No explanation is offered at this stage as to what has caused these transient dependencies, but similar application to stock² indices indicates that these results are

²The results are not shown, but are available from the authors upon request.

likely to be prevalent in many financial markets. A re-application of Engle's test for ARCH over shorter windows of data suggest that there are long periods of time where there is no significant evidence of ARCH in the data at all.

Our results may have fundamental consequences for financial econometrics, for these transient epochs of dependence could not be generated by any kind of ARCH or GARCH model, so that no amount of tinkering with the specification could capture these important features of the data. Recent modifications, such as to augment the GARCH equation with structural breaks in the mean, as suggested by Booth *et al.* (1994) or to generalize the model to allow the parameters of the GARCH equation to be drawn from one of several regimes (Hamilton and Susmel, 1994), still could not capture these dependencies. It appears, then, that the search for a convenient and parsimonious representation of financial asset returns which is also capable of modelling these properties must continue.

REFERENCES

- Bollerslev, T. (1986) Generalised autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–27.
- Booth, G.G., Martikainen, T., Sarkar, S.K., Virtanen, I. and Yli-Olli, P. (1994) Nonlinear dependence in Finnish stock returns, *European Journal of Operational Research*, **74**, 273–83.
- Box, G.E.P. and Pierce, D.A. (1970) Distributions of residual autocorrelations in autoregressive integrated moving average models, *Journal of the American Statistical Association*, **65**, 1509–26.
- Brock, W.A., Dechert, W.D. and Scheinkman, J.A. (1987) A test for independence based on the correlation dimension, mimeo Department of Economics, University of Wisconsin at Madison.
- Brock, W.A., Hsieh, D.A. and LeBaron, B. (1991) *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*, M.I.T. Press, Reading, Mass.
- Brooks, C. (1996) Testing for nonlinearities in daily sterling exchange rates, *Applied Financial Economics*, **6**, 307–17.
- Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987–1007.
- Engle, R.F., Hendry, D.F. and Trumble, D. (1985) Small sample properties of ARCH estimators and tests, *Canadian Journal of Economics*, **18**, 66–93.
- Hall, S.G., Miles, D.K. and Taylor, M.P. (1989) Modelling asset prices with time-varying betas *The Manchester School*, **57**(4), 340–56.
- Hamilton, J.D. and Susmel, R. (1994) Autoregressive conditional heteroskedasticity and changes in regime, *Journal of Econometrics*, **64**, 307–33.
- Hinich, M.J. (1995). Testing for dependence in the input to a linear time series model, *Journal of Nonparametric Statistics*, **6**, 205–21.
- Hinich, M.J. and Patterson, D.M. (1995) Detecting epochs of transient dependence in white noise, Mimeo., University of Texas at Austin.