Applied Economics Letters, 2011, **18**, 1321–1325

Routledge
Taylor & Francis Group

Stock returns in emerging markets and the use of GARCH models

Claudio A. Bonilla[a,*] and Jean Sepúlveda[b]

[a]*School of Economics and Business, University of Chile, Santiago, Chile*
[b]*School of Business and Economics, Universidad del Desarrollo, Ainavillo 456, Concepción, na Chile, Chile*

We use the Hinich portmanteau bicorrelation test to detect for the adequacy of using GARCH (Generalized Autoregressive Conditional Heteroscedasticity) as the data-generating process to model conditional volatility of stock market index rates of return in 13 emerging economies. We find that a GARCH formulation or any of its variants fail to provide an adequate characterization for the underlying process of the 13 emerging stock market indices. We also study whether there exist evidence of ARCH effects, over windows of 200, 400 and 800 observations, using Engle's LM (Lagrange Multiplier) test, and find that there exist long periods of time with no evidence of ARCH effects. The results suggest that policymakers should use caution when using autoregressive models for policy analysis and forecast because the inadequacy of GARCH models has strong implications for the pricing of stock index options, portfolio selection and risk management. Specially, measures of spillover effects and output volatility may not be accurate when using GARCH models to evaluate economic policy.

I. Introduction

Predicting returns on assets such as stocks and currencies is of crucial interest in empirical finance. Researchers and practitioners in finance are interested not only in predicting the mean of variables such as the aforementioned but also in forecasting the associated volatility of those means. Recent evidence shows that risk is time-varying, that is, there exist relative volatile periods that alternate with a more calmed ones.

Variance is a central parameter of any distribution. In option pricing the variance is the only unknown necessary to value the option (Black and Scholes, 1973). In the regulation of financial institutions, banking capital adequacy depends on the measurement of the Value-at-Risk that would depend on the volatility of returns of the portfolio of assets held by banks (Basel Committee, 2001). Following the banking regulation, other financial regulators have adopted Value-at-Risk as the appropriate risk measure to control insurance companies.

Most of the regulators have adopted the RiskMetrics methodology (Mina and Yi Xiao, 2001) to estimate volatility, which is an approximation based on an exponential smoother, the Exponentially Weighted Moving Average (EWMA), applied to the returns squared of a GARCH(1,1) (Generalized Autoregressive Conditional Heteroscedasticity) model.

Models of the GARCH type have spread through the finance industry, especially in regulation and volatility prediction. Thus, it is relevant to ask how appropriate the use of such models is.

Many models used in empirical finance to describe returns and volatility are linear. However, there are several indications that nonlinear models may be more appropriate. Franses and van Dijk (2002) enumerated some of the characteristic features of financial time series that suggest the necessity for considering nonlinear time-series models: (a) large returns (in absolute terms) occur more frequently than one might expect under the assumption that data are normally

*Corresponding author. E-mail: cbonilla@fen.uchile.cl

Applied Economics Letters ISSN 1350–4851 print/ISSN 1466–4291 online © 2011 Taylor & Francis
http://www.informaworld.com
DOI: 10.1080/13504851.2010.537615

distributed; (b) such large absolute returns tend to appear in clusters; (c) large negative returns appear more often than the large positive ones in stock markets, the opposite may occur with exchange rates; and (d) volatile periods are often preceded by large negative returns. Features (a) and (c) suggest the usefulness of models that have different regimes for returns. Features (b) and (d) suggest the relevance of models allowing for a description of time-varying volatility.

The most commonly used nonlinear model both in finance and in economics incorporates nonlinearity in the variance but not in the mean. These are models of (Generalized) Autoregressive Conditional Heteroscedasticity introduced by Engle (1982) (ARCH) and extended by Bollerslev (1986) (GARCH). GARCH models are popular because they are capable of describing not only the feature of volatility clustering but also certain other characteristics of financial time series, such as their pronounced excess kurtosis or fat-tailedness. Despite their broad adoption, 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 (Hall *et al.*, 1989).

The adequacy of using the ARCH/GARCH formulations is a question that has received less attention in the empirical finance literature. If the formulation commonly used in the analysis of the financial data is not adequate, then any policy recommendation or financial conclusion derived from the results could be misguided.

Our results on the test of GARCH inadequacy for 13 emerging stock markets are then important for policymakers in deciding the best model to forecast volatilities.

The structure of this article continues as follows. Section II provides a brief literature review of nonlinear behaviour in financial markets. Section III presents the data to be used in this study. Section IV briefly presents the Hinich portmanteau bicorrelation test and the Engle's LM test. Section V presents the empirical results obtained, and Section VI concludes.

II. Related Literature

Evidence of nonlinearities in stock market indices, exchange rates and labour markets data have been documented in recent literature that employ nonlinearity tests developed in the last two decades (Panagiotidis and Pelloni, 2003; Lim *et al.*, 2004; Bonilla *et al.*, 2006).

One of the questions that nonlinearity tests answered is the adequacy of ARCH/GARCH formulations in stock markets and that is the aim of this article, but with the novelty of using emerging stock market indices data. We concentrate on two aspects of GARCH models: the assumption of strict stationarity of GARCH models, and the presence of conditional heteroscedasticity over long sub-periods of the series. The first issue can be tested using an application of the Hinich third-order portmanteau test. The second aspect can be tested by using Engle's LM (Lagrange Multiplier) test. To our knowledge, this is the first time that the validity of the GARCH formulation will be formally analysed for these 13 emerging markets.

In the last few years, there have been several studies investigating the GARCH adequacy for modelling exchange rates and stock returns on different economies. Brooks and Hinich (1998) studied a set of 10 daily Sterling exchange rates and found that GARCH type of models cannot capture the statistical structure present in the data. Liew *et al.* (2003) and Lim *et al.* (2004) analysed several exchange rates for the Asian economies, finding that autoregressive models and GARCH models, respectively, are inadequate for capturing the behaviour of exchange rates in those countries. Bonilla *et al.* (2006) studied the main Latin American exchange rate returns series and found the inadequacy of modelling exchange rates, in those countries, by a GARCH formulation. Starica (2004) tested and rejected the hypothesis that a GARCH(1,1) process is the true data-generating process of the returns of the S&P 500 stock market index between 1957 and 2003. Brooks and Hinich (1998) reported several European stock indices and found that GARCH model cannot be considered a full representation of the process generating financial market returns. Lim *et al.* (2005) employed the Hinich portmanteau bicorrelation test to determine the adequacy of the GARCH model for eight Asian stock markets and found that this model cannot provide an adequate characterization for the underlying process of these Asian stock market indices.

Emerging market economies are an interesting subject. The political and financial instability that arises from time to time in most of these countries produces episodic nonlinearities in the stock markets indices (Bonilla *et al.*, 2006; Romero-Meza *et al.*, 2007). Besides, emerging markets are usually characterized by less developed stock markets and financial systems.

We will check whether the common failure of GARCH formulation encountered in the American, European and Asian stock markets is also a present characteristic of the 13 emerging economies as well.

III. The Data

Analysis is done using daily stock market index returns from the stock market indices of MERVAL (Argentina), BOVESPA (Brazil), IGPA (Chile), IGBC

Table 1. Number and percentage of significant windows of the binary transformed data for emerging stock market indices

Country	Index	Period	Number of observations	Number of windows	Number of significant windows	Percentage of significant windows (%)
Argentina	MERVAL	9/12/93–10/1/06	3000	120	6	5
Brazil	BOVESPA	28/4/93–11/1/06	3125	125	5	4
Chile	IGPA	3/1/90–10/1/06	3975	159	24	15
Colombia	IGBC	3/1/91–10/1/06	3400	136	24	18
Czech Republic	PX	5/1/95–24/8/10	3886	111	2	2
Hungary	BUX	5/1/95–24/8/10	3906	111	2	2
Indonesia	DJTM	2/1/92–18/5/10	4794	136	2	2
Mexico	IPC	11/11/91–10/1/06	3525	141	12	9
Peru	IGBVL	3/1/86–6/10/04	4500	180	26	14
Poland	WIRR	3/1/95–24/8/10	3907	111	4	4
Singapore	DTSINGL	2/1/92–18/5/10	4794	136	6	4
Thailand	DJTHAIL	3/1/92–18/5/10	4793	136	2	2
Venezuela	IBC	3/1/94–10/1/06	2900	116	4	3

(Colombia), PX (Czech Republic), BUX (Hungary), DJTM (Indonesia), IPC (Mexico), IGBVL (Peru), WIRR (Poland), DTSINGL (Singapore), DJTHAIL (Thailand) and IBC (Venezuela). Sample periods are indicated in Table 1. Data are split into a set of non-overlapping windows of 35 observations length (i.e. approximately 7 trading weeks). Daily returns are constructed as $r_t = \ln(p_t/p_{t-1})$, where p_t is the closing price of the market stock index in day t. This can be interpreted as a continuously compounded daily return.

IV. The Hinich Portmanteau Bicorrelation Test, the C-Test and the Engle's LM Test

We use the Hinich portmanteau bicorrelation test statistic (denoted as H-statistic) and the simple autocorrelation C-statistic with the windowed test procedure. Due to space constraints, we refer the interested reader to Hinich (1996) and Hinich and Patterson (2005) for a mathematical derivation of these statistics and their small sample properties.

We also study the parameter stability of GARCH models and the transient nature of ARCH effects. Engle (1982) developed a test for conditional heteroscedasticity in the context of ARCH models based on the LM principle. First, we run a linear regression against a constant and save the residuals $\hat{\varepsilon}_t$. Then, we square the residuals and regress them on 1 own lag to test for ARCH of order 1. The LM test statistics is defined as TR^2 (the number of observations multiplied by the coefficient of multiple correlation) from the last regression. This statistic is distributed as a χ^2 (1) under the null of no order 1 ARCH effects.

This test is common in the empirical finance literature, but we use a slightly different approach. We apply the test over a set of relatively short nonoverlapping windows of length 200, 400 and 800 observations with the aim of discovering whether there is strong evidence of ARCH over all time periods, or just for short periods of time.

V. Empirical Results

We define a 1% nominal threshold for the p-values of the Hinich portmanteau test. This means that we would expect to have a 1% of the nonoverlapped windows significant only by chance. However, our analysis shows that this is not the case. Table 1 presents the number and percentage of significant windows for the binary transformed data.

Results show that in the 13 economies, a larger number of windows are significant than the 1% threshold level. Thus, data are unlikely to be generated by a stationary GARCH model. This result provides evidence of the inadequacy of using GARCH models for the 13 stock market indices analysed. Interestingly, the higher percentage of significant windows corresponds to three Latin American countries (Chile, Colombia and Peru). These three countries have in common a relatively small stock market, in terms of number of companies and daily transactions, compared with the other countries in the sample.

These results are similar to the previous studies that analysed the inadequacy of using GARCH type of models in Asian countries (Lim *et al.*, 2005) and European countries (Brooks and Hinich, 1998).

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

Country	Index	Length	Number (percentage) of nonsignificant windows at		
			10%	5%	1%
Argentina	MERVAL	200	7(47%)	10(67%)	10(67%)
		400	1(14%)	2(14%)	3(14%)
		800	0(0%)	0(0%)	0(0%)
Brazil	BOVESPA	200	7(44%)	8(50%)	12(75%)
		400	2(25%)	2(25%)	4(50%)
		800	0(0%)	1(25%)	1(25%)
Chile	IGPA	200	3(15%)	3(15%)	5(25%)
		400	0(0%)	0(0%)	1(10%)
		800	0(0%)	0(0%)	0(0%)
Colombia	IGBC	200	8(47%)	8(47%)	9(53%)
		400	2(25%)	2(25%)	3(38%)
		800	0(0%)	0(0%)	1(25%)
Czech Republic	PX	200	12(63%)	13(68%)	15(79%)
		400	3(33%)	4(44%)	6(67%)
		800	0(0%)	0(0%)	1(4%)
Hungary	BUX	200	9(47%)	11(58%)	12(63%)
		400	3(33%)	3(33%)	4(44%)
		800	0(0%)	0(0%)	0(0%)
Indonesia	DJTM	200	12(52%)	12(52%)	15(65%)
		400	3(27%)	6(54%)	7(64%)
		800	2(40%)	2(40%)	2(40%)
Mexico	IPC	200	9(50%)	11(61%)	12(67%)
		400	3(33%)	3(33%)	3(33%)
		800	0(0%)	1(25%)	1(25%)
Peru	IGBVL	200	5(23%)	5(23%)	7(32%)
		400	0(0%)	1(9%)	1(9%)
		800	1(20%)	1(20%)	1(20%)
Poland	WIRR	200	7(37%)	10(53%)	13(68%)
		400	1(11%)	2(22%)	2(22%)
		800	0(0%)	0(0%)	1(25%)
Singapore	DTSINGL	200	13(57%)	14(61%)	19(83%)
		400	2(18%)	4(36%)	6(55%)
		800	0(0%)	0(0%)	0(0%)
Thailand	DJTHAIL	200	12(52%)	14(61%)	16(70%)
		400	3(27%)	3(27%)	3(27%)
		800	0(0%)	0(0%)	1(20%)
Venezuela	IBC	200	6(43%)	9(64%)	9(64%)
		400	2(29%)	2(29%)	2(29%)
		800	1(25%)	1(25%)	1(25%)

Table 2 shows the number of nonsignificant windows using a sub-sample size of 200, 400 and 800 observations at the 10%, 5% and 1% significance level. When data are divided into sub-samples, we find the existence of long periods of time with no evidence of ARCH effects. We can observe that in most cases (countries and length of the windows) the percentage of nonsignificant windows is different from zero.

Yet, these results cannot be attributed to a fall in the power of the test at smaller samples. For example, Engle *et al.* (1985) showed that 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.

VI. Conclusions

In this study, we research for the adequacy of using GARCH type formulations for modelling stock market indices in 13 emerging economies: 7 from Latin America, 3 from East Europe and 3 from Asia. To that end we use the Hinich portmanteau bicorrelation test and the Engle's ARCH test. These results indicate that the GARCH formulation fails to capture the statistical structure of the stock market returns series for all the countries studied here. In this sense, the GARCH formulation fails to capture the time-varying nature of market returns analysed because it treats coefficients as fixed and being drawn from only one regime. For instance, Bonilla *et al*. (2006) showed that Latin American stock market indices are characterized by transient epochs of dependence surrounded by long periods of white noise.

The Engle's test for ARCH effects reaffirms the results. Indeed, for both short and long periods of time we find a significant number of windows with no evidence of ARCH effects.

Although the number of significant windows we find is smaller than those found by Lim *et al*. (2005), for Asian stock markets, it is still high and important. We hypothesize that political instability in most of the analysed countries and the relative low importance of these economies in the world context translate into inefficient stock markets.

We conclude by pointing out that policymakers should use caution when using autoregressive models for policy analysis and prediction because the inadequacy of GARCH models has strong implications for the pricing of stock index options, portfolio selection and risk management.

Our results are in line with previous literature that studied the behaviour of stock market indices of American, European and Asian countries.

References

Basel Committee on Bank Supervision (2001) *The New Basel Capital Accord*, Bank for International Settlements, Basel.

Black, F. and Scholes, M. (1973) The pricing of options and corporate liabilities, *Journal of Political Economy*, **81**, 637–54.

Bollerslev, T. (1986) Generalised autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–27.

Bonilla, C., Romero-Meza, R. and Hinich, M. J. (2006) Episodic nonlinearities in the Latin American stock market indices, *Applied Economics Letters*, **13**, 195–9.

Brooks, C. and Hinich, M. (1998) Episodic nonstationarity in exchange rates, *Applied Economics Letters*, **5**, 719–22.

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.

Franses, P. H. and van Dijk, D. (2002) *Non-Linear Time Series Models in Empirical Finance*, Cambridge University Press, Cambridge.

Hall, S. G., Miles, D. and Taylor, M. (1989) Modelling asset prices with time-varying betas, *The Manchester School*, **57**, 340–56.

Hinich, M. (1996) 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. (2005) Detecting epochs of transient dependence in white noise, in *Money, Measurement and Computation*, (Eds.) M. Belongia and J. Binner, Palgrave, London.

Liew, V., Chong, T. and Lim, K. (2003) The inadequacy of linear autoregressive models for real exchange rates: empirical evidence from Asian economies, *Applied Economics*, **35**, 1387–92.

Lim, K. P., Hinich, M. J. and Liew, V. (2004) Adequacy of GARCH models for ASEAN exchange rates return series, *International Journal of Business and Society*, **5**, 17–32.

Lim, K. P., Hinich, M. J. and Liew, V. (2005) Statistical inadequacy of GARCH models for Asian stock markets: evidence and implications, *International Journal of Emerging Market Finance*, **4**, 263–79.

Mina, J. and Yi Xiao, J. (2001) Return to RiskMetrics: the evolution of a standard. RiskMetrics Group, Inc. Available at www.riskmetrics.com (accessed 4 December 2010).

Panagiotidis, T. and Pelloni, G. (2003) Testing for nonlinearity in labour markets: the case of Germany and the UK, *Journal of Policy Modeling*, **25**, 275–86.

Romero-Meza, R., Bonilla, C. and Hinich, M. J. (2007) Nonlinear event detection in the Chilean stock market, *Applied Economics Letters*, **14**, 987–91.

Starica, C. (2004) Is GARCH(1,1) as good a model as the Nobel prize accolades would imply?, Economics Working Paper No. 0411015, Archive Econ WPA, University of Gothenburg, Sweden.