COMPROBANTE DE RECEPCIÓN DE INFORME FINAL

Nº PROYECTO: 11060447    ETAPA: 2007
INVESTIGADOR RESPONSABLE: CLAUDIO BONILLA MELENDEZ
RUT: 12266242-K

TÍTULO: MARKET EFFICIENCY: STUDYING NONLINEAR BEHAVIOUR IN EMERGING BONDS SPREADS AND CHECKING FOR THE ADEQUACY OF GARCH FORMULATIONS

DISCIPLINA: FINANZAS(CS. ECONOM/ADMI)

GARCH Inadequacy for Emerging Market Bonds Spreads: Empirical Evidence from Latin America

Claudio Bonilla*
Universidad del Desarrollo

Abstract

In this paper we analyze the adequacy of using GARCH as the data generating process to model conditional volatility for Latin-American emerging markets bonds spreads series. Using the Hinich portmanteau bicorrelation test, we find that a GARCH formulation or any of its variants fail to provide an adequate characterization for the underlying process of the main Latin American bonds spreads. Policy makers need to be careful when using autoregressive models for policy analysis and forecast because the inadequacy of GARCH models has strong implications for the pricing of emerging markets bonds, portfolio selection, and risk management. In particular, measures of spillover effects and output volatility may not be correct when using GARCH type of models to evaluate economic policy.

JEL Classification: C12, G19
Keywords: GARCH models, Hinich portmanteau bicorrelation test, emerging markets bonds spreads

* The author acknowledge the financial support of FONDECYT project 11060447
I. Introduction

Forecasting future returns on assets such as stocks, bonds and currencies is of primary interest in empirical finance. Since the trade-off between risk and return plays a prominent role in many financial models, and that volatility is often regarded as a risk measure, researchers and practitioners in finance are interested not only in accurate forecast of returns, but also in forecasts of the associated volatility. Recent evidence shows that risk is time-varying, that is, there exists relative volatile periods that alternate with a more calmed ones.

Volatility is a central parameter of any distribution. In option pricing is the only unknown necessary to value the option (Black and Scholes, 1973). In financial institutions regulation, the banking capital adequacy depend on the measurement of the Value-at-Risk which would depend on the volatility of returns of the portfolio of asset 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. For the case of Chile see SVS (2002, 2004).

Most of the regulators have adopted the RiskMetrics methodology (Mina and Yi, 2001) to estimate volatility, which is an approximation based on an exponential
smoother, the exponentially weighted moving average (EWMA), applied to returns squared of a model GARCH(1,1).

Given the current use of GARCH type of model in regulation and volatility forecasting, for policy modeling is relevant to ask how sensible the use of ARCH-GARCH type of 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) enumerate 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 distributed, (b) such large absolute returns tend to appear in clusters, (c) large negative returns appear more often than large positive ones in stock markets and bonds spreads, 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.
Hsieh (1989) provides a useful division of nonlinear dependencies into two broad categories – additive and multiplicative nonlinearity. Additive nonlinearity, or nonlinearity-in-mean, enters a process through its expected value. With multiplicative nonlinearity, or nonlinearity-in-variance, each element can be expressed as the product of a zero-mean random element and a nonlinear function of past elements, so that the nonlinearity affects the process through its variance.

An example of additive nonlinear model is the bilinear model (see e.g., Granger and Anderson, 1978; Subba Rao and Gabr, 1980). For this type of model the unconditional mean is equal to zero, and its realization would not exhibit any autocorrelation. However, in contrast to a normal linear process, such a process would exhibit nonzero bivariate correlations. One way to test for the existence of nonzero bivariate correlations is through the use of the Hinich's H test that we review below (see Hinich and Patterson, 2005; Hinich, 1996). However, the most commonly used nonlinear model in finance and economics exhibits nonlinearity in variance, rather than in the mean. These are models of (Generalized) Autoregressive Conditional Heteroscedasticity introduced by Engle, 1982 (ARCH) and extended by Bollerslev, 1986 (GARCH). GARCH models enjoy such popularity 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).

A question, however, that has received less attention in the empirical finance literature is the adequacy of these ARCH/GARCH formulations. 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 can be misleading.

One of the most successful econometrics models to estimate conditional variance applied to financial and economic data are the ARCH-GARCH family. For example, Brissimis and Chionis (2004) use a GARCH specification to explore the effects of foreign exchange market intervention by the European Central Bank and the Bank of Japan upon the conditional variance of the euro-dollar and euro-yen exchange rates. Also, emerging markets bonds spreads have been widely studied empirically using GARCH family of models (see Edwards and Sumsel, 2003; Ciferalli and Paladina; 2004 or Jaque and Rojas 2003 for example). In the same way, several researchers have investigated on the propagation of the fluctuations of stock return on economic activity. For example, Hassapis and Kalyritis (2002) examine
empirically the response of output growth to shock in real stock returns. They find that a portion of excess output volatility in OECD countries can be explained by innovations in foreign real returns. They conclude that the global economic growth is more exposed to financial instability originating from abroad. These links among financial markets of different countries stress the importance for modeling volatility properly. Our results on the test of GARCH inadequacy for Latin American bonds spreads are then important for policy makers in deciding the best model to foresee volatilities.

The structure of the remainder of the paper is as follows. The next section provides a brief literature review of nonlinear behavior in financial markets. Section III sketches the GARCH model. Section IV introduces and explores the data to be used in this study. Section V presents the Hinich portmanteau bicorrelation test and the Engle’s LM test. Section VI presents the empirical results obtained. Final conclusions are given in section VII.

II. Related Literature
Evidence of nonlinearities in stock market indices, exchange rates, bonds spreads and labour markets data have been documented in recent literature that employ nonlinearity tests developed in the last two decades. (Bonilla et al. 2006, Lim et al. 2004, Bonilla et al., 2008 and Panagiotidis and Pelloni, 2003) One of the questions that nonlinearity tests answered is the adequacy of ARCH/GARCH formulations in exchange rates and stock markets, but our research has the novelty of using Latin American bonds spreads data. We concentrate on two aspects of GARCH models. First, the assumption of strict stationarity of GARCH models can be tested using an application of the Hinich third order portmanteau test to the main Latin American bonds spread data. Second, the presence of conditional heteroscedasticity over long sub-periods of the series, using Engle’s LM test. To our knowledge, this is the first time that the validity of the GARCH formulation will be formally analyzed for the main Latin American bonds spreads.

In the last few years, there have been several interesting studies on Latin American financial data. However, the methodology and the focus used on them are different than ours. Among them, we have Fujii (2005), Fernandez (2005), Verma (2005) and Fernandez-Serrano and Sosvilla-Rivero (2003). Fujii (2005) examines the causal linkage among several emerging stock markets in Asia and Latin America since 1990. Fernandez (2005) focuses on return spillovers in stock markets at different

We found previous literature that check for ARCH/GARCH adequacy of exchange rates in European, Asian and Latin American markets. Brooks and Hinich (1998) studied a set of ten daily Sterling exchange rates. Liew et al. (2003) and Lim et al. (2004) analyzed several exchange rates for the Asian economies. Bonilla et al. (2007) studied the main Latin American exchange rate returns series. Results of these empirical papers indicate that the usual GARCH formulation used for modeling the behavior of exchange rates fails to capture the data generating process of these series for these economies. These results are consistent with recent findings in the literature that provides evidence in favor of nonlinear structures in the exchange rates series (see for example Ma and Kanas, 2000; Sarno, 2000 or Sarantis, 1999).
There are also some studies that investigate the issue of ARCH/GARCH adequacy for stock market indices in the North American, European and Asian markets. Starica (2004) tests and rejects 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. They found that this kind of model cannot provide an adequate characterization for the underlying process of these Asian stock market indices. Thus, testing the main Latin American bonds spreads would complement the empirical evidence of GARCH adequacy as an appropriate data generating process.

Latin American economies are an interesting subject. The political and financial instability that arises from time to time in these countries produce episodic nonlinearities in the stock markets indices (Bonilla et al. 2006 and Romero-Meza et al. 2007). We will check if the common failure of GARCH formulation encountered in the American, European and Asian stock markets is also a present characteristic of the Latin American economies as well.
In this study we analyze if major Latin American bonds spreads are correctly modeled using the GARCH formulation. If not, any policy and financial conclusion derived from previous studies using autoregressive conditional heteroscedastic models are potentially misleading.

III. GARCH Models

The empirical finding that financial time series present volatility clustering effects, and that volatility occurs in bursts – that is, after a long period of calm, a period of rising volatility arise -, makes highly unlikely to find constant variance across time in financial series. To parameterize this fact, researchers make use of a conditional variance model, where the variance of the errors is allowed to change over time in an autoregressive conditional heteroscedastic framework (Engle, 1982 and Bollerslev, 1986)

An observed time series $y_t$ can be written as the sum of a predictable and an unpredictable part,

$$y_t = E[y_t | \Psi_{t-1}] + \varepsilon_t \quad (1.1)$$
Where $\Psi_{t-1}$ is the information set consisting of all relevant information up to and including time $t-1$. These types of models allow the conditional variance of $\{\varepsilon_t\}$ to vary over time. A convenient representation of this is

$$
\varepsilon_t = z_t h_t^{1/2}
$$

$$
\varepsilon_t \mid \Psi_{t-1} \sim N(0, h_t) \quad (1.2)
$$

$$
h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}
$$

Where $z_t$ is independent and identically distributed with zero mean and unit variance. This GARCH formulation captures the fact that volatility is changing over time. The variation corresponds to a weighted average among long term average variance, volatility in the previous period, and the fitted variance of previous period as well.

Higher order GARCH formulation are not usually necessary in finance because current variance ($h_t$) implies infinitely long memory of past innovation. Therefore previous formulation is a standard model used to parameterize any financial time series (see Bollerslev et. al., 1992 for an account of the use of the use of ARCH models in finance).
Ours results show that the usual GARCH formulation used for stock market returns behavior modelling fails to capture the data generating process of the main Latin American stock market indices. A re-application of Engle's test for ARCH over shorter windows of data suggest that long periods of time exist with no significant evidence of ARCH in the data at all.

IV. The Data

Analysis is done using daily bonds spread data obtained from JP Morgan securities for seven Latin American economies: Argentina, Brazil, Chile, Colombia, Mexico, Peru and Venezuela. Sample periods for the indices are: Argentina from December 1993 to January 2006, Brazil from April 1993 to January 2006, Chile from January 1990 to January 2006, Colombia from January 1991 to January 2006, Mexico from November 1991 to January 2006, Peru from January 1986 to October 2004, and Venezuela from January 1994 to January 2006. Data are split into a set of non-overlapping windows of 25 observations length (i.e. approximately five trading weeks).
V. 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. Let the sequence $\{Z(t)\}$ denote the standardized sampled data process (created by subtracting the sample mean of the window, and dividing by its standard deviation), where time unit $t$ is an integer. The test procedure employs non-overlapped data window, thus if $n$ is the window length, then the $k$-th window is $\{Z(t), Z(t+1), \ldots, Z(t+n-1)\}$. The next non-overlapped window is $\{Z(t+1), Z(t+1+1), \ldots, Z(t+n+1)\}$, where $t_{k+1} = t_k + n$. The null hypothesis for each window is that $Z(t)$ are realizations of a stationary pure noise process that has zero bicorrelation. The alternative hypothesis is that the process generated within the window is random with some non-zero bicorrelations $C_{ZZZ}(r,s) = E[Z(t)Z(t+r)Z(t+s)]$ in the set $0 < r < s < L$, where $L$ is the number of lags that define the window. For a mathematical derivation of this statistics and its small sample properties the interested reader is referred to Hinich (1996). We thus state without derivation of the statistics test, denoted $H$ and $C$ respectively.
\[ H = \sum_{s=2}^{L} \sum_{r=1}^{L-1} \left[ G^2(r,s)/(T-s) \right] \sim \chi^2((L-1)L/2) \] (1.3)

where
\[ G(r,s) = \sum_{i=1}^{T} \left[ Z(t_i) Z(t_{i+r}) Z(t_{i+s}) \right] \]

and
\[ C = \sum_{i=1}^{T} \left[ C^2(r)/(T-r-1) \right] \sim \chi^2(L) \] (1.4)

where
\[ C(r) = \sum_{i=1}^{T} Z(t_i) Z(t_{i+r}) \]

The number of lags \( L \) is specified as \( L = n^b \) with \( 0 < b < 0.5 \), where \( b \) is a parameter under the choice of the analyst. Based on results from Monte Carlo simulations (see Hinich and Patterson, 2005) it is recommended to use \( b = 0.4 \) in order to maximize the power of the test while ensuring a valid approximation to the asymptotic theory. In this test procedure, a window is significant if the \( H \) or \( C \) statistic rejects the null of pure noise at the specified threshold level.

It is possible to use the above test to check whether a GARCH formulation represent an adequate characterization of the data. This is achieved by
transforming the returns into a set of binary data denoted \( \{x_p(t)\} \) where \( x_p(t) = 1 \) if \( Z_p(t) > U \) and \( x_p(t) = -1 \) if \( Z_p(t) < U \). If the original \( \{Z_p(t)\} \) are generated by a ARCH or GARCH process, then \( \{x_p(t)\} \) will be a stationary pure noise series provided that innovations have a symmetric density with zero mean. The binary transformed data has moments which are well-behaved with respect to the asymptotic theory (see Hinich, 1996).

If the number of windows of binary transformed rates which have significant \( C \) or \( H \) statistic rejecting the null of whiteness at a specified threshold level for the \( p \)-value is much larger than \( p \), then the original process is unlikely to be generated by a GARCH process. This rejection might be due to serial dependence in the innovations but this violates a critical assumption for ARCH and GARCH models. If innovations are dependent, then the statistical properties of the parameter estimates are unknown.

Also we study 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 Lagrange Multiplier (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, and is distributed as a $\chi^2(1)$ under the null of no order 1 ARCH effects. We run the Engle's LM test using the software Eviews 5.0.

This test is very common in the empirical finance literature, but we use a slightly different approach. We apply the test over a set of relatively short non-overlapping 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.

VI. Empirical Results

We define a 0.1% nominal threshold for the p-values of the Hinich Portmanteau test. This means that we would expect to have a 0.1% of the non-overlapped windows significant only by chance. However our analysis shows a complete different result. Table 1 presents the number and percentage of significant windows for the binary transformed data.
Results show that a larger number of windows are significant than the 0.1% 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 Latin American Bonds spreads. This is important because most of the literature that analyze bonds spreads uses GARCH/ARCH generating processes to study economic aspects of them like contagion in emerging markets, co-movements in bonds markets, risk premia in bonds spreads and other related issues.

Table 2 shows the number of non-significant windows using a sub-sample size of 200, 400, and 800 observations at 10%, 5%, and 1% significance level. When data are split into sub-samples, we find the existence of long periods of time with 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 sub-samples are aggregated.
Yet, these results cannot be attributed to a fall in the power of the test at smaller samples. For example, Engle et al. (1985) show 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.

VII. Conclusions

In this paper we use the Hinich portmanteau bicorrelation test to check for the adequacy of using a GARCH formulation for modeling the Latin American bonds spreads. Our results indicate that the GARCH formulation fails to capture the data generating process of the bonds spreads series for all Latin American markets studied. Policy makers need to be careful when using autoregressive models for policy analysis and forecast because the inadequacy of GARCH models has strong implications for the pricing of emerging markets bonds, portfolio selection, and risk management. Policy conclusions derived from the use of GARCH models to study financial propagation of exchange and financial markets intervention and output volatility must be careful considered. Our results are consistent with previous related literature that analyzed stock market indices and exchange rates behavior of american, european and asian countries.
In particular, the GARCH model fails to capture the time-varying nature of the spreads, and treats coefficients as fixed and being drawn from only one regime. Latin American bonds spreads have been shown to be characterized by transient epochs of dependence surrounded by long periods of white noise. A reapplication of Engle’s test for ARCH over shorter windows of data suggests that there are long periods of time where there is no significant evidence of ARCH in the data at all.

Our results present a high percentage of significant windows. Our hypothesis is that political instability and low relative importance of these economies in the world context reflects into inefficient financial markets. A further investigation that looks at political and economic events leading to potential non-linearities and chaos within the significant windows for each country needs to be done in the future. Since politics affects the economy (Brooks et al. 2000 and Romero-Meza et al. 2007) and that this region recurrently has political crises, we should expect to find a considerable amount of non-overlapped windows significant.

References


Starica, C. (2004) Is GARCH(1,1) as good a model as the Nobel prize accolades would imply?, Economics Working Paper Archive EconWPA.


### Table 1
Number and Percentage of Significant Windows of the Binary Transformed Data for Latin American Bonds Spreads

<table>
<thead>
<tr>
<th>Country</th>
<th>Period</th>
<th>Number of observations</th>
<th>Number of windows</th>
<th>Number of significant windows</th>
<th>Percentage of sig. windows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>12/9/93 - 1/10/06</td>
<td>3000</td>
<td>120</td>
<td>6</td>
<td>5%</td>
</tr>
<tr>
<td>Brazil</td>
<td>4/28/93 - 1/11/06</td>
<td>3125</td>
<td>125</td>
<td>5</td>
<td>4%</td>
</tr>
<tr>
<td>Chile</td>
<td>1/3/90 - 1/10/06</td>
<td>3975</td>
<td>159</td>
<td>24</td>
<td>15%</td>
</tr>
<tr>
<td>Colombia</td>
<td>1/3/91 - 1/10/06</td>
<td>3400</td>
<td>136</td>
<td>24</td>
<td>18%</td>
</tr>
<tr>
<td>Mexico</td>
<td>11/11/91 - 1/10/06</td>
<td>3525</td>
<td>141</td>
<td>12</td>
<td>9%</td>
</tr>
<tr>
<td>Peru</td>
<td>1/3/86 - 10/6/04</td>
<td>4500</td>
<td>180</td>
<td>26</td>
<td>14%</td>
</tr>
<tr>
<td>Venezuela</td>
<td>1/3/94 - 1/10/06</td>
<td>2900</td>
<td>116</td>
<td>4</td>
<td>3%</td>
</tr>
</tbody>
</table>

### Table 2
The number of non-significant windows using Engle’s test for ARCH

<table>
<thead>
<tr>
<th>Country</th>
<th>Length</th>
<th>Number (percentage) of non-significant windows at</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10%</td>
<td>5%</td>
</tr>
<tr>
<td>Argentina</td>
<td>200</td>
<td>7 (47%)</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>1 (14%)</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Brazil</td>
<td>200</td>
<td>7 (44%)</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>2 (25%)</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Chile</td>
<td>200</td>
<td>3 (15%)</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>0 (0%)</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Colombia</td>
<td>200</td>
<td>8 (47%)</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>2 (25%)</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Mexico</td>
<td>200</td>
<td>9 (50%)</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>3 (33%)</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Peru</td>
<td>200</td>
<td>5 (23%)</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>0 (0%)</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>1 (10%)</td>
</tr>
<tr>
<td>Venezuela</td>
<td>200</td>
<td>6 (43%)</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>2 (29%)</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>1 (25%)</td>
</tr>
</tbody>
</table>