Testing GARCH and RV Exchange Rate Volatility Models using Hinich Tricorrelations

Sanja Dudukovic[1]
Franklin University – Switzerland

Abstract: The aim of this paper is to enlighten a need to test two most popular volatility models: the GARCH-ARMA model based on a daily returns and the RV – ARMA model based on 30 min intraday HF data, in terms of non Gaussian Time Series Analysis. The ability of the models to perform digital whitening and to produce "white" innovations is tested on seven exchange rates, including JPY/EUR, USD/EUR, CAD/USD, CHF/EUR, CHF/USD, USD/GBP and GBP/EUR, the daily data as well as 30 min data.In the first step, stationary ARMA-GARCH models of different orders were built and the best model was chosen by using AIC and Box-Pierce test based on the innovations of daily squared returns. In the second step, realized daily volatilities, defined as the sum of intraday 30 min squared returns, are used to estimate the RV-ARMA volatility model parameter and to calculate forecasting errors. In the third step, fourth order cumulants are calculated for 20 lags for all currencies and used to perform the Hinich test. Finally, it was shown that whitening of squared returns (GARCH) and daily realized volatilities are not efficient in either case. The finding of serial dependence in innovations which can be categorized as a deep structure phenomenon opens the question if it is still appropriate to simply equate the presence of non Gaussianity and nonlinearity with the presence of outliers. Further improvement is to be achieved by estimating the model parameters by using Higher Order Cumulant function prior to performing the HOC based testing.

Key Words: Volatility Forecasting, , Higher Order Cumulant Function , GARCH Model, ARMA Model, Exchange Rate Volatility , Model testing , Hinich test.

JEL Classification Numbers: G15, G17

INTRODUCTION

Eversince the GARCH volatility framework was established a cohesive body of GARCH literature has encapsulated many of the aspects of its ability to capture the market stylized facts of a foreign exchange rate market (FX). While numerous studies have compared the forecasting abilities of the historical variance and GARCH models, no clear winner has emerged. In a scrupulous review of 93 such studies, Poon and Granger (2003) reported that 22 find that historical volatility forecasts future volatility better out-of-sample, while 17 studies find that GARCH models forecast better. Indeed, there are many variants of ARCH/ GARCH models which are developed to improve the out-of-sample volatility forecasting performance.
Indeed, there are many variants of ARCH/GARCH models which are developed to improve the out-of-sample volatility forecasting performance. These models have many strong proponents, who believe that GARCH models are currently the best obtainable forecast estimators. However, most of the empirical studies on the subject in recent years have found

[1] E mail : Sdudukovic@fc.edu

ISBN : 9780974211428

no clear-cut results in improving forecasting performances of this class of GARCH models; Carrol & Kearney [12], for example Brooks, Burke and Persand (2001) used DJ composite daily data to test in- and out-of-sample forecasts obtained with GARCH, EGARCH, GRJ and HS (historical volatility) models. The coefficient of determination (R^2) achieved was around 25% for each of the models.

Contrary to the classic setting for economic forecast evaluation, the volatility is not directly observable, but rather intrinsically latent. Accordingly, any ex post assessment of forecast precision must account for a fundamental errors-in-variable problem coupled with the realization-measurement of the predicted variable.

The availability of the high-frequency intraday data has led a number of recent studies to promote and evaluate the use of so-called realized volatilities (RV), constructed from the summation of finely sampled squared high-frequency returns, as a practical method for improving the ex post volatility measures. Assuming that the sampling frequency of the squared returns utilized in the realized volatility computations approaches zero, the realized volatility then consistently estimates the true (latent) integrated volatility.

Unluckily, market microstructure frictions deform the measurement of returns at the highest frequencies so that, e.g., tick-by-tick return processes obviously violate the theoretical semi-martingale restrictions implied by the no-arbitrage assumptions in continuous-time asset pricing models. Thus realized volatility measures constructed directly from the ultra high-frequency returns appear to be biased. As such, the integrated volatility is regularly measured with error (Bai, Russell, and Tiao (2000)).

Traditional methods of comparing volatility models insofar have been: Mean Forecast Error (MSE) produced by those models and its many variants; maximum likelihood value and AIC or BIC criteria. Most recently the other approach is taken. That is to say, given a set of characteristic features or exchange rate stylized facts such as volatility clustering, fat tail phenomena, leverage effect or Taylor effect, one may ask the following question: "Have popular volatility models been parameterized in such a way that they can accommodate and explain the most common stylized facts visible in the data?" Models for which the answer is positive may be viewed as suitable for practical use. For example, Teräsvirta (1996) investigated the ability of the GARCH model to reproduce series with high kurtosis and, at the same time, positive but low and slowly decreasing autocorrelations (AC) of squared observations. Carnero, Peña & Ruiz (2004) compared the ARSV model and the GARCH model using the kurtosis autocorrelation relationship in squared returns as their benchmark. Bai, Russell & Tiao (2003) also compared GARCH and ARSV models in terms of kurtosis and AC.

Ultimately, non Gaussian Time Series Analysis (TSA) is gaining new importance in the context of volatility modeling and risk management. Ideally, in terms of TSA, a good volatility model should have a capacity to perform “digital whitening“ of stock market squared returns and therefore to produce white innovations (iid), which are known as forecasting errors or simply as driving noise. The aim of this paper is to test the ability of two best known volatility models, GARCH and RV, to produce non correlated and independent innovations. The organization of the paper is as follows. The GARCH and the RV models are defined in Section 2; The Box& Pierce test and the Hinich tricorrelation test are introduced in Section 3.The same section presents the introduction to higher order moments and cumulants. The data description and model building results are presented in Section 3. Section 4 presents comparative innovation analysis and model testing results. Section 5 contains conclusions and suggestions for further research.

 ISBN : 9780974211428

2. VOLATILITY MODELS

The fact that stock market returns are often characterized by volatility clustering – which means that periods of a high volatility are followed by periods of a high volatility and periods of a low volatility are followed by periods of a low volatility – implies that the past volatility could be used as a predictor of the volatility in the next. Although the autocorrelation of the returns is insignificant at all frequencies, the autocorrelations of the squared absolute returns persist within a very long time interval demonstrating a long memory in volatility.

The kurtosis of the returns is much higher than that of a normal distribution at intraday frequency and tends to decrease as the return length increases. Thus return probability density functions (pdfs) are leptokurtic with a fat tails. It was believed that both volatility clustering and fat tails could be explained by using the well-known GARCH model. That is to say, those stylized facts are not seen only in autocorrelation function, kurtosis and skewness of squared returns, but also in the third and the fourth order cumulant functions.

2.1 The GARCH model

Let e_t denote a discrete time stationary stochastic process. The GARCH (p, q) process is given by the following set of equations:

$$r_t = \log(P_t) - \log(P_{t-1}) \quad (1)$$

$$r_t = x(_k)g(_k) + e_t \quad (2)$$

$$e_t = v_t\sqrt{h_t}$$

$$e_t/_{t-1} \approx N(0, h_t) \quad (3)$$

$$h_t = \alpha_0 + \sum_1^p \alpha_i e^2_{t-i} + \sum_1^q \beta_j h_{t-j} \quad (4)$$

where p_t represents stock prices; e_t represents random returns; $x(_k)$ is a vector of explanatory variables; g $(_k)$ is a vector of multiple regression parameters; h_t is the conditional volatility; α_i is autoregressive; and β_j is the moving average parameter as related to the squared stock market index residuals. An equivalent ARMA representation of the GARCH (p, q) model (Bollerslev, 1982, pp. 42-56) is given by:

$$e_t^2 = \alpha_0 + \sum_1^p(\alpha_i + \beta_i)e^2_{t-i} + \nu_t - \sum_1^q \beta_j \nu_{t-j} \quad (5)$$

where $\nu_t = e_t^2 - h_t$ and, by definition, it has the characteristics of (i.i.d) white noise. h_t is known as GARCH variance.

In this context, the GARCH (p, q) volatility model is simply an Autoregressive Moving Average, ARMA (p,q) model in e_t^2 driven by i.i.d noise ν_t, which is Gaussian random variable. It is worth stressing that the GARCH variance, h_t, in time series analysis, appears to be merely an estimate of the squared de-trended SM returns e_t^2.

The best known ARMA model building methodology is known to be the Box-Jenkins (B-J) iterative methodology, which includes three steps: model order determination, parameter estimation and model testing (Box and Jenkins, 1970).

 ISBN : 9780974211428

The B-J methodology assumes that each stationary time series can be treated as an output from the AR(p), MA(q) or ARMA (p,q) filter, which has as an uncorrelated and Gaussian innovations, known as "white noise" $\{\nu_\tau\}$.
The ARMA model has the following form: $A(Z)\, e_t^2 = B(Z) * \nu_t$, where **Z** is a backward shift operator:
$e_{t-1}^2 = Z^{-1} e_t^2 : e_{t-k}^2 = Z^{-k} e_t^2$, and where $A(Z) = 1 - \alpha_1 Z^{-1} - \alpha_2 Z^{-2} - \alpha_p Z^{-p}$ and
$B(Z) = 1 - \beta_1 Z^{-1} - \beta_2 Z^{-2} - \beta_q Z^{-q}$ are characteristic transfer functions of orders p and q respectively The roots of the characteristic functions of the ARMA model must be within the unit cycle to guarantee stationarity and invertibility of the model.

2.2.Realized Volatility models

Recently Corsi (2009) introduced an alternative approach to construct an observable proxy for the latent volatility by using intraday high frequency data.
His work was inspired by Merton (1980), who showed that the integrated volatility of a Brownian motion can be approximated to an arbitrary precision using the sum of intraday squared returns.

$$IV_t = \int_{t-1}^{t} \sigma^2(s)ds \quad (6)$$

So, in this integrated framework, the Integrated Variance (IV) is considered to be the population measure of actual return variance. Namely, it was proved that the sum of intraday squared returns converges (as the maximal length of returns go to zero) to the integrated volatility of the returns making it possible to construct an error free estimate of the actual volatility. This nonparametric volatility estimator is known as realized volatility (RV).

$$RV_t = \sum_1^{\infty} r_{t,i}^2 \quad (7)$$

Taken correctly, this theory suggests that one should sample prices as often as possible. This would direct to estimate IV_t by RV_t from tick-by-tick data. However, as was noted in Merton (1980) “in practice, the choice of an even-shorter observation interval introduces another type of error which will swamp the benefit long before the continuous limit is reached”. The modern terminology for this phenomenon is known as market microstructure effects that cause the observed market price to diverge from the efficient price. All said, market structure effect introduces a bias that grows as the sampling frequency increases. This motivated the idea of viewing the observed prices, p_t, as noisy measures of the latent true price.

Indeed, in practice, empirical data at very small time (RV) make a strongly biased estimator in case of small SM return interval. Therefore, a trade-off arises: on one hand, statistical theory would impose a very high number of return observations to reduce the stochastic error of the measurement; on the other hand, market microstructure comes into play, introducing a bias that grows as the sampling frequency increases. Given such a trade-off between measurement error and bias, a simple way out is to choose, for each financial variable, the shortest return interval at which the resulting volatility is still not significantly affected by the bias. This approach has been pursued by Andersen et al. (2001) and (Mauller at al.(1997), who agree on a return interval of 30 minutes for the most highly liquid exchange rates leading to only 48 observations per day.

 ISBN : 9780974211428

Literally, it is believed that the realised volatility, defined as the sum of intraday, 30 min squared returns, provides a more accurate estimate of the latent volatility than the estimate based on daily squared returns.

$$RV_t=\sum_1^{48} r^2_{t,i} \qquad (8)$$

The theoretical and empirical properties of realized volatility are derived in (Andersen, Bollerslev, Diebold and Labys, 2001) for foreign exchange. They found that realised volatility distribution is nearly Gaussian. Further empirical evidence is provided in (Andersen, Bollerslev, Diebold & Ebens, 2001) for U.S. equities. In this article the ARMA model applied to RV is tested:

$$RV_t = \alpha_0 + \sum_1^p \alpha_i RV_{t-i} + u_t - \sum_1^q \beta_j u_{t-j} \qquad (9)$$

Estimated realized volatility is than calculated by using the formula: $ERV = RV_t - u_t$

3. MODEL TESTING METHODS

There are two tests which can be applied to test the null hypothesis that the ARMA model innovation time series represent a white noise. The first is the well known Box& Pierce test which can be applied if innovations – driving noise is independent and identically distributed Gaussian process. In the case of non Gaussian probability density function, Box-Pierce test would not show model inadequacy since it is based only on second order statistics, which is no longer sufficient for parameter estimation.
All stationary time series are time reversible (TR) but the contrary is not true. Visually, TIR demonstrate a tendency of a variable to rise rapidly to local maxima and then to decay slowly. This time reversibility amounts to temporal symmetry in the probabilistic structure of the process and is typical for stock market variables. TR cannot be evaluated by using the second order cumulants – autocorrelation function. Therefore the test based on higher order cumulant function is more appropriate in non Gaussian case.

3.1 Box-Pierce Q test

As for diagnostic checking, if obtained model is appropriate and the parameter estimates are consistent and efficient, for the particular time series, then the model innovations νt would be uncorrelated random deviates, and their first L sample autocorrelations:

$$AC(k) = \sum \nu(t)\nu(t-k) / \sum \nu_t^2$$

would have a multivariate normal distribution, Box-Pierce (1970). They showed also that the AC(k), k=0, 1, 2…L, are uncorrelated with variances which could be approximated by: $V(AC(k) = (n-k)/n(n+2) \approx 1/n$, from which it follows specifically that the statistic $n(n+2)\sum(n-k)^{-}AC(k)^2$ would for large n be distributed as $\chi 2$ with L degrees of freedom; or as a further approximation,

$$n\sum (AC(k)^2 \approx \chi_L^2$$

 ISBN : 9780974211428

When applied to the ARMA parameter estimation, degree of freedom must be changed to L-p-q, where p and q are the orders of the autoregressive and moving average operators.

3.2 The Hinich Test

As found empirically, in the case of exchange rate and stock market returns, driving noise is not Gaussian Subsequently the second order moment and correlation function do not represent "sufficient statistics", nether for the ARMA parameter estimation, nor for the model testing. In fact, it is well known that for a non-Gaussian process, the higher order moments exist and are different from zero. Basic cumulants are defined somewhat abstrusely as follow.

3.2.1. Cumulants

The rth moment of a real-valued random variable ***X*** with density f(x) is

$$\mu_r = E(X^r) = \int_{-\infty}^{\infty} x^r f(x)\, dx$$

for integer r = 0, 1,…mother value is assumed to be finite. Provided that it has a Taylor expansion about the origin, the moment generating function:

$$\mu_r = E(X^r) = \int_{-\infty}^{\infty} x^r f(x)\, dx$$

is an easy way to combine all of rth moments into a single expression. The rth moment is the rth derivative of M at the origin.

$$K(\xi) = \log M(\xi) = \sum_r \kappa_r \xi^r / r!.$$

The cumulants K_r are the coefficients in the Taylor expansion of the cumulant generating function about the origin

$$K(\xi) = \log M(\xi) = \sum_r \kappa_r \xi^r / r!.$$

Evidently m_o = 1 implies k_0=0. The relationship between the first few moments and cuinulants, obtained by extracting coefficients from the expansion, is as follows:

 ISBN : 9780974211428

$$\kappa_1 = \mu_1$$
$$\kappa_2 = \mu_2 - \mu_1^2$$
$$\kappa_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3$$
$$\kappa_4 = \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4.$$

Cumulants of order r > 2 are called semi-invariant on account of their behavior under affine transformation of variables (Thiele 1898). This behavior is considerably simpler than that of moments. However, moments about the mean are also semi-invariant, so this property alone does not explain why cumulants are useful for statistical purposes. The term cumulant was coined by Fisher (1929) on account of their behavior under addition of random variables. Let $S = X + Y$ be the sum of two independent random variables. The moment generating function of the sum is the product

$$M_S(\xi) = M_X(\xi)M_Y(\xi),$$

and the cumulant generating function is the sum:

$$M_S(\xi) = M_X(\xi)M_Y(\xi),$$

Consequently, the rth cumulant of the sum is the sum of the rth cumulants. By extension, if $X_1,...X_n$ are independent and identically distributed, the rth cumulant of the sum is $n\kappa^r$ and the rth cumulant of the standardized sum $n^{-1/2}(X1+X2\ X_n)$ is $n^{1-r/2}k^r$. Provided that the cumulants are finite, all cumulants of order r > 3 of the standardized sum tend to zero. As an example of the way these formulae may be used, let X be a scalar random variable with cumulants k1, k2, k3, k4… By translating the second formula in the preceding list, we find that the variance of the squared variable

$$\mathrm{var}(X^2) = \kappa_4 + 4\kappa_3\kappa_1 + 2\kappa_2^2 + 4\kappa_2\kappa_1^2,$$

reducing to $\kappa^4 + 2k_2^2$ if the mean is zero.

In the area of digital signal processing, Giannakis (1990) was the first to show that the third and the fourth order cummulant functions can be efficiently estimated as following:

$$C^3_r(\tau1,\tau2)= (\sum(r(t)r(t+\tau_1)r(t+\tau_2))/n, \qquad \tau2=1,2..L\ ,\ \tau1=1,2...L \qquad (10)$$

$$C^4(\tau1,\tau2,\tau3,)= (\sum(r(t)r(t+\tau_1)r(t+\tau_2)\ r(t+\tau_3))/n - -C^2_r(\tau1)\ C_r(\tau2-\tau3) - C^2_r(\tau2)\ C_r(\tau3-\tau1)-C^2_r(\tau3)\ C_r(\tau1-\tau2), \qquad (11)$$

where n is a number of observations and where the second-order cumulant $C^2_r(\tau)$ is just the autocorrelation function of the time series of returns $r_{t,}$ t=1,2.3…n.

The test statistics suggested by Hinich (1996) is the portmanteau tricorrelation test statistics, denoted as the G statistics. This statistics, which is a third-order extension of the standard Box-Pierce correlation test for white noise, tests the null hypothesis that the ARMA

 ISBN : 9780974211428

model standardized innovations are realizations of a pure white noise process that has zero auto correlations and tricorrelations. The analytical form of the Hinich test is the following:

$$C(\tau_{1,}\tau_{2,\,,}\tau_{3,})=(N-s)^{-1/2}\,(\sum(\nu(t)\nu\,(t+\tau_1)\nu\,(t+\tau_2)\,(t+\tau_3) \qquad (11)$$

Hinich proved the theorem which states that his test statistics H_N, follows χ^2 distribution with M degrees of freedom M=L*(L-1)*L/3, where L is a number of lags. Hinich (1996) statistics is presented as the normalized sum of values $G^2(\tau_{1,}\tau_2)$ for $(1<\tau_{1<}\tau_2<L$:

$$G=(N-s).^{1/2}\ C(\tau_1,\tau_2,\tau_3)$$

$$H_N=\sum_{m=3}^{L-1}\ \sum_{s=2}^{m-1}\sum_{1}^{s-1}[G^2(\tau_{1,}\tau_{2,}\tau_3)\ ]\ , \qquad (2)$$

The distribution of G is approximately chi-squared with L(L - 1)(L-2)/3 degrees of freedom for large N if $L = T^c$, $(0 < c<.5)$ under the null hypothesis that the observed process is pure white noise (iid). The parameter c is chosen by the user. Thus, under the pure white noise, null hypothesis, $U = F(H_N)$ has a Gaussian (0,1) distribution, where F is the cumulative distribution function of a chi-squared distribution. In principle, the test can be applied to either the source returns or to the model residuals.

4. EMPIRICAL ANALYSIS

4.1 GARCH-ARMA results

The ARMA-GARCH empirical analysis is based on daily quotations of closing daily exchange rates for the period from Sep 21, 2012 to March 20, 2013, taken from Bloomberg. The common sample of exchange rate description is presented in Table 1.

Table 1. Descriptive statistics of the squared FX daily returns

	R2JPYEUR	R2USDEUR	R2CADUSD	R2CHFEUR	R2CHFUSD	R2USDGBP
Mean	0.15	0.04	0.02	0.01	0.03	0.03
Median	0.05	0.02	0.01	0.00	0.01	0.01
Maximum	1.36	0.47	0.15	0.28	0.28	0.29
Minimum	0.00	0.00	0.00	0.00	0.00	0.00
Std. Dev.	0.25	0.06	0.03	0.03	0.05	0.05
Skewness	3.12	4.06	2.45	5.25	2.62	2.81
Kurtosis	13.50	26.38	8.87	35.47	10.11	12.57
Jarque-Bera	677.65	2782.05	265.30	5288.11	353.76	559.04

The reported statistics confirm the skewed distributions across all currencies. In addition, the sample kurtosis for each currency is well above the normal value of 3. Jarque-Bera values show that all FX return distributions are leptokurtic and depart significantly from Gaussian distribution.

The ARMA-GARCH parameter estimates based on OLS method are given in Table 2. The table presents only the best stationary model for each currency and is chosen when achieving the minimum Akaike Information Criterion (AIC).

The Box &Pierce (1980) test of the null hypothesis that the first K autocorrelations of covariance stationary innovations are zero, in the presence of statistical dependence, was performed. The results are in Table 2.

Table2. ARMA-GARCH parameter estimates.

Currency	C	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)	R2	AIC	Q
JPYEUR	0.159	0.240	-0.163	0.471	-0.231	0.544	-0.226	0.232	-0.693	0.450	-0.407	0.222	-0.115	36.166
st.error	0.059	0.306	0.259	0.148	0.218	0.243	0.337	0.281	0.123	0.292	0.329			
USDEUR*		-1.877	-1.844	-0.846	-0.037		0.968	0.025	-0.938	-0.951		0.575	-2.781	25.418
		0.094	0.180	0.180	0.093		0.023	0.018	0.021	0.024				
CADUSD	0.010	0.498	-0.879	0.880	-0.449	0.924	-0.476	0.965	-0.952	0.470	-0.970	0.124	-3.327	31.493
	0.022	0.027	0.031	0.023	0.028	0.028	0.022	0.017	0.014	0.021	0.016			
CHFEUR		0.281	0.370	-0.317	-0.110	0.699	-0.125	-0.416	0.603	0.174	-0.795	0.222	-4.803	26.896
		0.077	0.092	0.080	0.078	0.065	0.075	0.081	0.045	0.070	0.067			
CHFUSD*		-1.130	-0.875	-1.051	-0.974	-0.124	0.161	-0.288	0.176	-0.037	-0.871	0.528	-2.397	32.374
		0.101	0.130	0.114	0.087	0.067	0.078	0.075	0.094	0.072	0.072			
USDGBP*		0.831	-0.938	0.193	0.043		-1.959	2.017	-1.406	0.407		0.572	-3.271	25.653
		0.161	0.138	0.143	0.083		0.171	0.265	0.261	0.143				

*denotes an ARMA model
** denotes a second difference model--i.e. d(RJPYEUR, 2)

The presented results show the best ARMA model for each currency in terms of AIC criterion. All presented models are stationary. Stationarity is achieved by taking the first or the second difference of FX returns. ARCH-ARMA parameter estimates show unexpectedly high coefficient of determination. Q statistics shows that residuals are non-correlated. The statistical properties of GARCH innovations are given in Table 3.

Table 3. ARMA-GARCH innovation statistical description

	RESR2JPYE	RESR2USD	RESR2CAD	RESR2CHF	RESR2CHF	RESR2USD	RESR2GBPEUR
Mean	0.0057	0.0023	-0.0027	0.0030	-0.0032	0.0051	0.0093
Median	-0.0534	-0.0080	-0.0122	-0.0005	-0.0180	-0.0071	-0.0050
Maximum	0.9920	0.3978	0.1251	0.2714	0.2550	0.2368	0.4167
Minimum	-0.4392	-0.0892	-0.0363	-0.0418	-0.0730	-0.0760	-0.2269
Std. Dev.	0.2149	0.0557	0.0311	0.0313	0.0547	0.0448	0.0785
Skewness	2.0918	3.8307	2.0813	6.1100	2.6506	2.2811	2.1602
Kurtosis	9.8540	25.6201	7.6652	51.4944	11.0768	10.4969	12.2986
Jarque-Bera	292.8428	2590.4050	177.5413	11358.8500	423.9068	349.7843	477.4624

As it can be seen from the table, kurtosis is extremely high for all currencies, which suggests a strong departure from the model assumption, which stated that GARCH residuals were supposed to be normally distributed.

4.2 RV-ARMA results

High frequency squared returns were used to create daily realized volatility for all currencies.

$RV_t=\sum_1^{48} r^2_{t,i}$

 ISBN : 9780974211428

RV data are presented in Figure 2.Their statistical properties are given in Table 4.

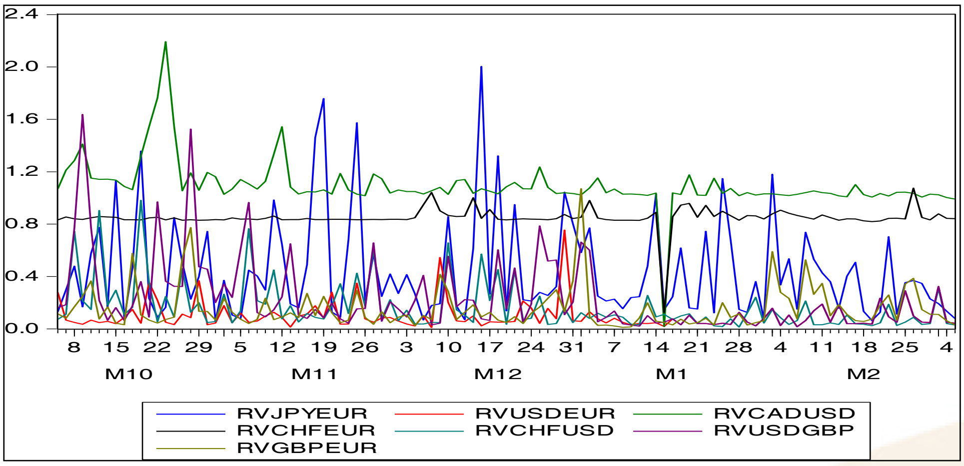

Figure 2. Daily realized volatility – all currencies

Table 4. Statistical description of daily realized volatilities

	RVJPYEUR	RVUSDEUR	RVCADUSD	RVCHFEUR	RVCHFUSD	RVUSDGBP	RVGBPEUR
Mean	0.46	0.10	1.10	0.85	0.17	0.24	0.15
Median	0.28	0.06	1.05	0.84	0.11	0.14	0.09
Maximum	3.37	0.75	2.19	1.07	0.98	1.64	1.07
Minimum	0.06	0.01	0.02	0.12	0.01	0.02	0.01
Std. Dev.	0.47	0.11	0.19	0.08	0.19	0.28	0.16
Skewness	2.90	3.06	0.78	-6.14	2.26	2.44	2.86
Kurtosis	15.35	15.09	21.32	62.19	8.40	10.36	13.70
Jarque-Bera	907.8	896.1	1648.0	17815.6	241.5	380.2	717.2

Table 4 clearly shows again departure from the Gaussian distribution. These realised volatilities are then used to make an RV-ARMA (p,q) model. The model parameters based on E-views software are presented in Table 5.

Table 5: RV-ARMA parameters

	C	AR(1)	AR(2)	AR(3	AR(4)	AR(5)	AR(6)	AR(7)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)	R^2	AIC:	Q:
JPYUR**		-1.53	-1.75	-1.70	-1.39	-1.00	-0.60	-0.25						0.74	1.90	24.03
		0.09	0.16	0.21	0.23	0.21	0.16	0.09								
USDEUR	0.10	0.40	0.26	-0.41	-0.34				-0.59	-0.24	0.62	0.51		0.29	-1.88	20.01
	0.01	1.06	1.11	0.36	0.46				1.04	1.31	0.53	0.62				
CADUSD	0.88	-0.05	0.67	0.43	-0.30	0.23			0.45	-0.59	-0.77	0.03	-0.10	0.35	-0.41	217.36
	0.16	0.64	0.33	0.47	0.54	0.16			0.64	0.58	0.33	0.71	0.24			
CHFEUR	0.85	-0.35	0.91	-0.06	-0.88	-0.12			0.44	-0.91	0.00	1.03	0.22	0.08	-2.56	23.01
	0.01	0.25	0.05	0.25	0.09	0.22			0.26	0.04	0.26	0.08	0.25			
CHFUSD*		-0.75	-0.98	-0.69	-0.95	-0.09			-0.29	0.27	-0.23	0.34	-0.95	0.60	-2.37	108.99
		0.07	0.05	0.07	0.05	0.07			0.02	0.02	0.02	0.02	0.02			
USDGBP*		-0.23	-1.36	-0.26	-0.91	-0.12			-0.86	1.30	-1.22	0.91	-0.92	0.62	-3.24	25.41
		0.07	0.02	0.10	0.02	0.07			0.03	0.01	0.04	0.02	0.03			
GBPEUR		-0.28	-0.29	-0.50	-0.63	0.16			-0.49	-0.10	0.22	0.28	-0.87	0.46	-0.81	
		0.21	0.18	0.16	0.19	0.12			0.19	0.25	0.23	0.22	0.17			

*denotes an ARIMA model
** denotes a second difference model--i.e. d(rvjapeur,2)

 ISBN : 9780974211428

As it can be seen from Table 5, Box-Pierce test statistics Q, applied to model residuals, shows that for two currencies hypothesis of non correlated innovations cannot be rejected. This finding suggests a question of the validity of the assumption that RV residuals are produced by a non correlated process.

4.3 COMPARATIVE ANALYSIS

While numerous studies have compared the forecasting abilities of the historical variance and GARCH models, no clear winner has emerged.
This paper is based on innovation analysis in terms of Hinich test. Innovations are calculated for both types of models: GARCH-ARMA and RV-ARMA. Figure 1 presents the GARCH-ARMA innovations for all seven currencies. Figure 2 presents RV-ARMA innovations. Proceeding with the data description, statistical properties of the GARCH-ARMA residuals and RV-ARMA innovations are presented in Table 6.1 and Table 6.2 respectively. From statistical description, it is obvious that neither of the volatility models has captured high kurtosis of squared daily returns or realized daily volatility in the case of seven currencies being tested.

Table 6.1. ARMA-GARCH innovation statistical description

	RESR2JPYEU	RESR2USDEU	RESR2CADU:	RESR2CHFEU	RESR2CHFUS	RESR2USDGI	RESR2GBPEUR
Mean	0.01	0.00	0.00	0.00	0.00	0.01	0.01
Median	-0.05	-0.01	-0.01	0.00	-0.02	-0.01	0.00
Maximum	0.99	0.40	0.13	0.27	0.26	0.24	0.42
Minimum	-0.44	-0.09	-0.04	-0.04	-0.07	-0.08	-0.23
Std. Dev.	0.21	0.06	0.03	0.03	0.05	0.04	0.08
Skewness	2.09	3.83	2.08	6.11	2.65	2.28	2.16
Kurtosis	9.85	25.62	7.67	51.49	11.08	10.50	12.30
Jarque-Bera	292.84	2590.41	177.54	11358.85	423.91	349.78	477.46

Table 6.2 RV-ARMA innovations description

	RESRVJPYEUR	RESRVUSDEUR	RESRVCADUSD	RESRVCHFEUR	RESRVCHFUSD	RESRVUSDGB	RESRVGBPEUR
Mean	0.00	0.01	-0.01	-0.02	-0.02	0.07	0.01
Median	-0.03	0.00	0.00	-0.05	-0.06	-0.07	-0.03
Maximum	0.66	0.76	0.21	0.45	1.14	3.15	0.79
Minimum	-0.11	-1.01	-0.71	-0.33	-0.47	-0.64	-0.27
Std. Dev.	0.11	0.16	0.08	0.14	0.24	0.52	0.15
Skewness	3.26	-1.14	-6.05	1.21	1.63	2.48	2.05
Kurtosis	17.75	23.03	58.53	5.03	8.37	13.61	9.74
Jarque-Bera	1180.31	1845.56	14669.51	45.08	179.59	622.56	282.62

This contradicts the finding by (Andersen, Bollerslev, Diebold and Labys, 2000) that residuals are nearly Gaussian.

ISBN : 9780974211428

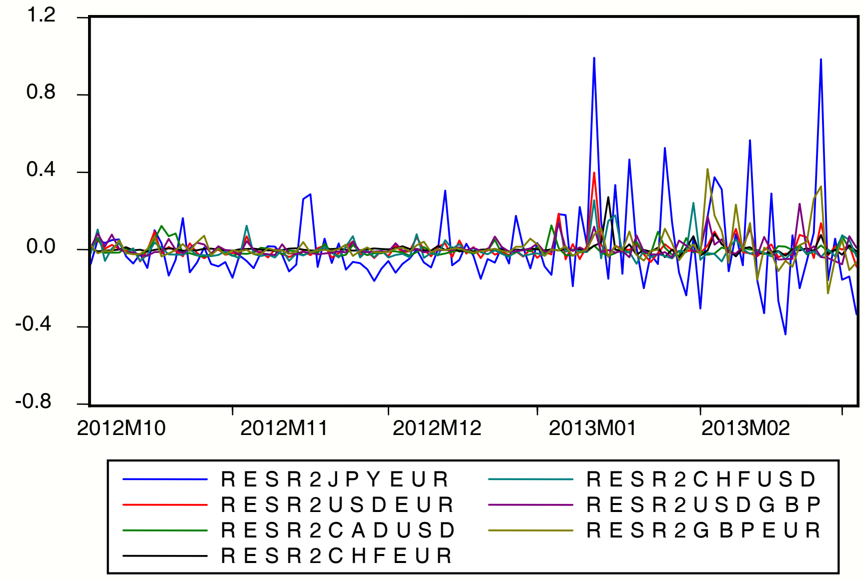

Figure 3 Daily GARCH-ARMA residuals

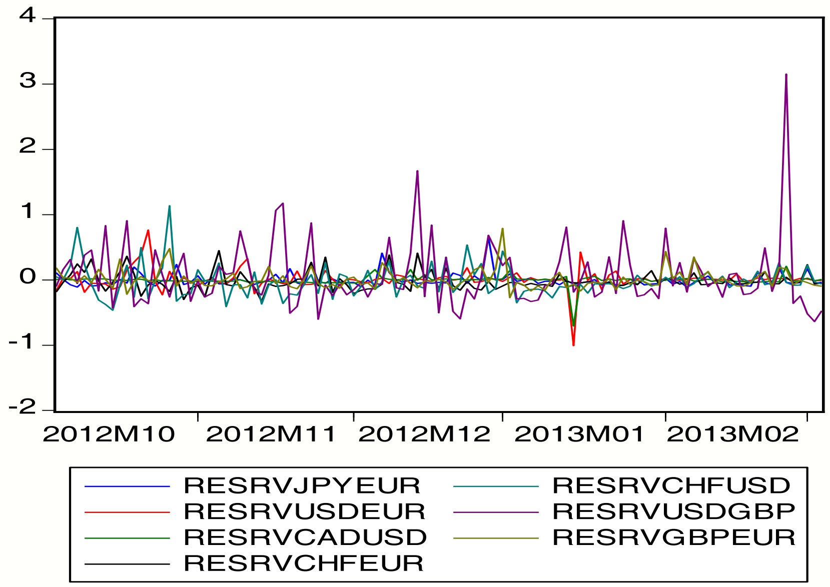

Figure 5 RV-ARMA residuals

The forth order cumulants are calculated according to equation (5) and are presented in Figures 6.1, 6.2, 6.3 and 6.4. These figures confirm, what is in line with the above considerations, that both ARMA-GARCH and ARMA-RV produce non Gaussian innovations. The null hypothesis that innovation - driving noise is IID process is evaluated tested using Hinch test. The results are presented in Table 7.

Table 7. Hinich test results for ARMA-RV and ARMA-GARCH innovations

The results show that in the case of 5 models, the null hypothesis which states that innovations are “white”, cannot be rejected. But in all other cases Hinich test was smaller that χ^2 critical (77.78), confirming that white model innovations are not produced on a regular basis.

 ISBN : 9780974211428

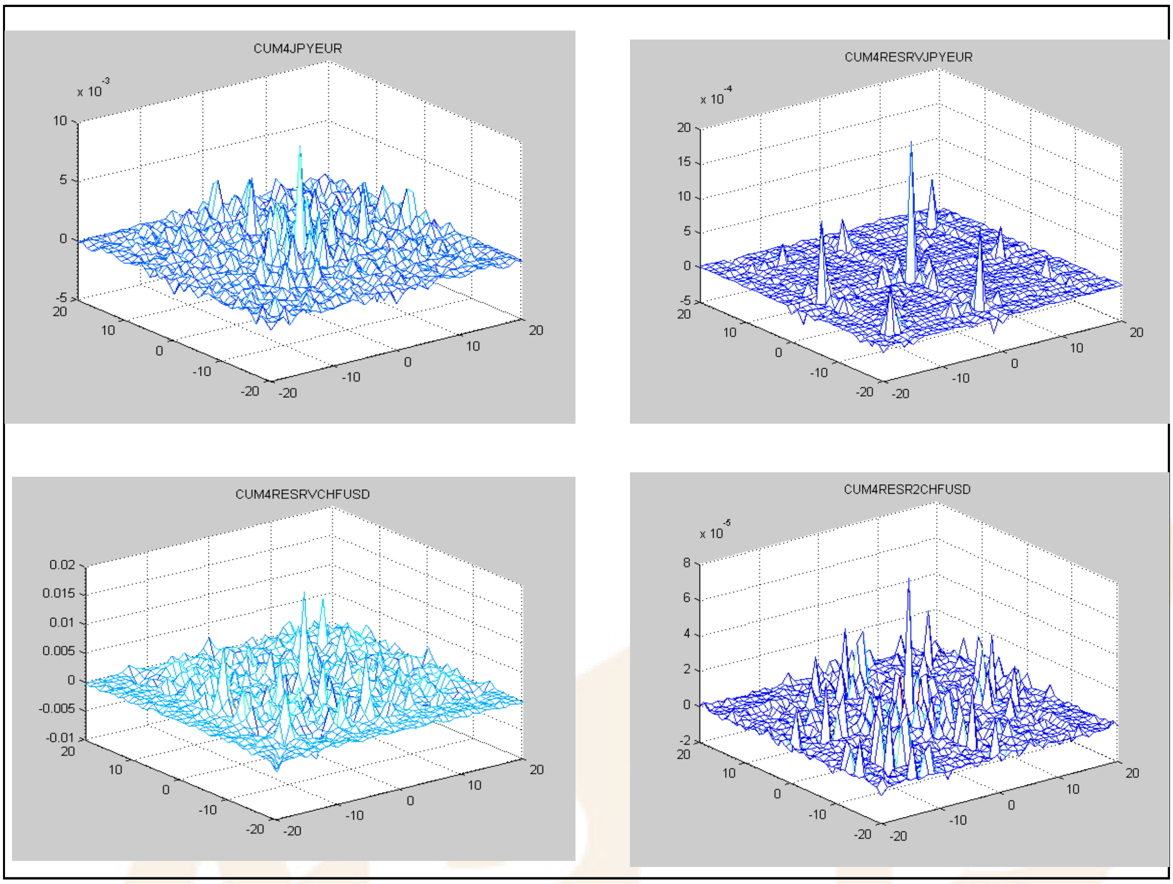

Figure 6.1. Fourth Order Cumulants – JPYEUR and CHFUSD exchange rates.

ISBN : 9780974211428

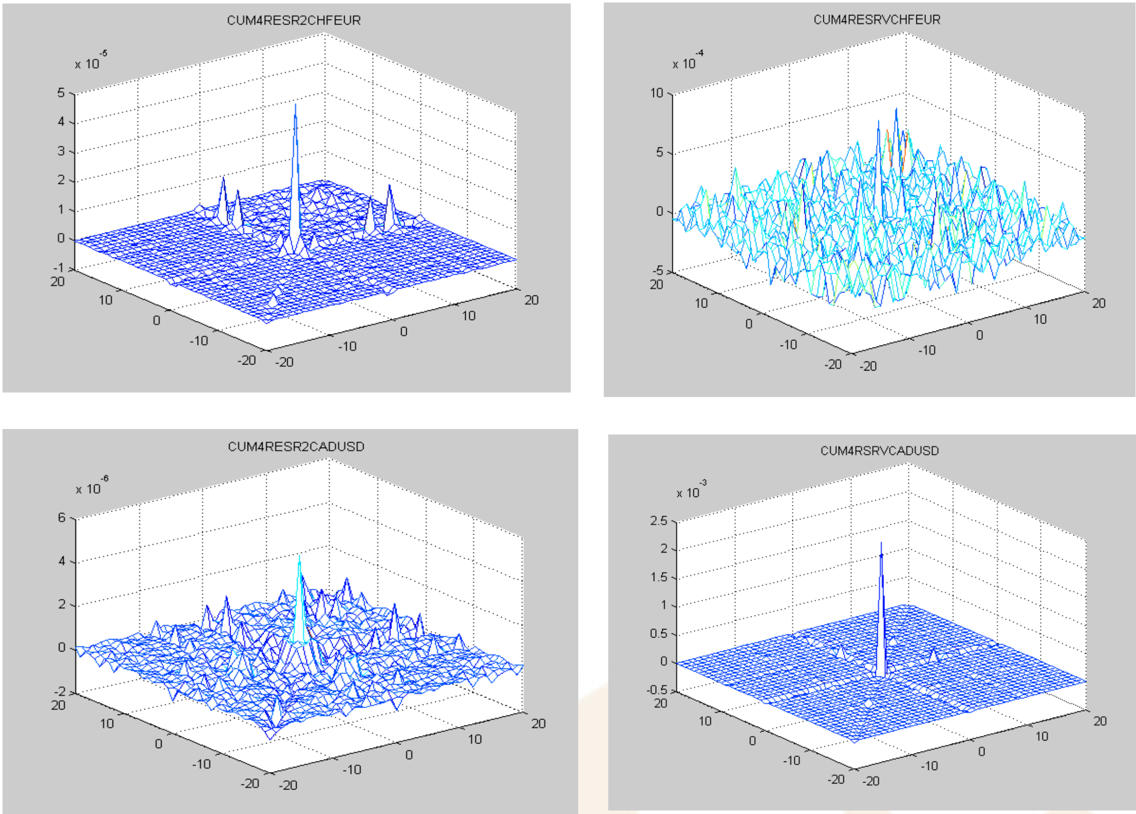

Figure 6.2 Fourth Order Cumulants – CHFEUR and CADUSD exchange rates.

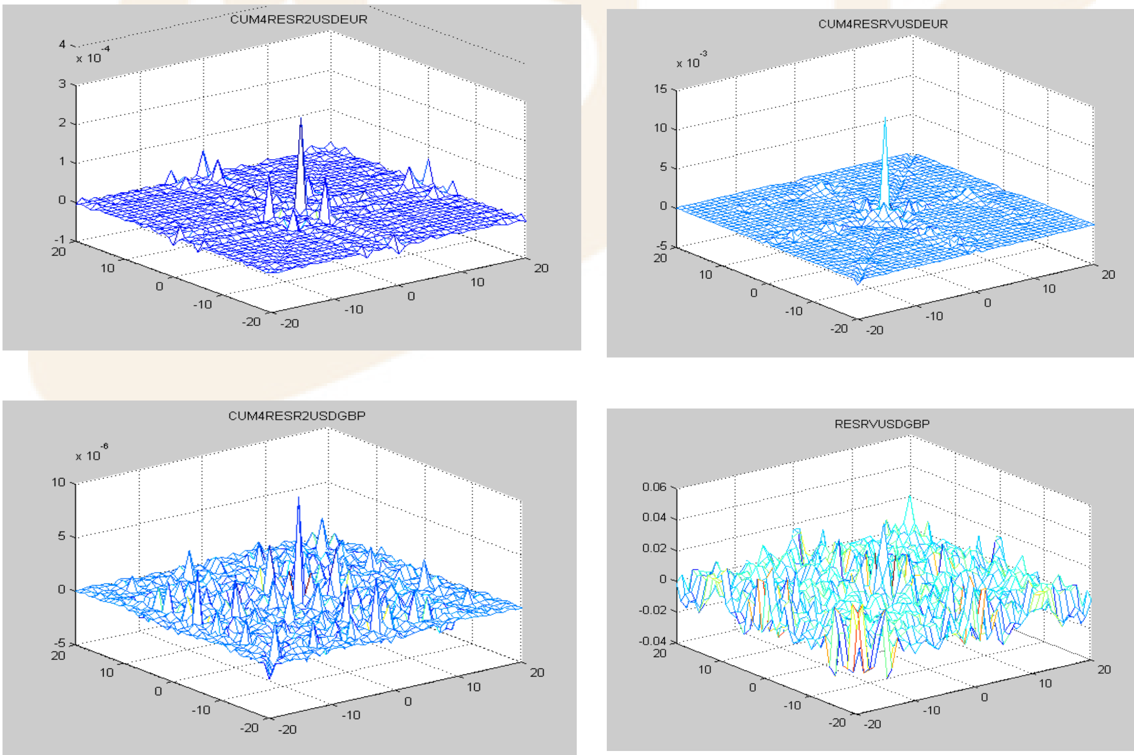

Figure 6.3. Fourth Order Cumulants – USDEUR and USDGBP exchange rates

 ISBN : 9780974211428

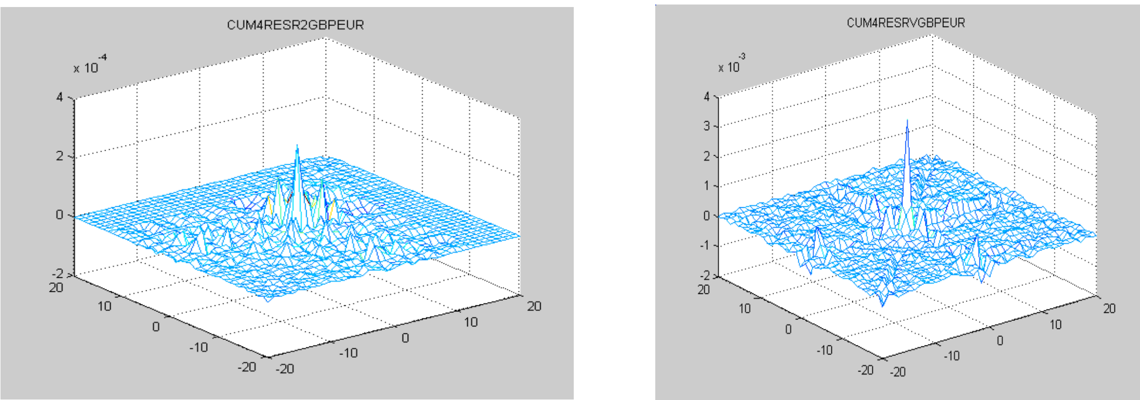

Figure 6.4. Fourth Order Cumulants – GBPEUR and GBPEUR exchange rates

CONCLUSION

This paper aimed to compare ARMA-GARCH and ARMA-RV volatility models in terms of the statistical properties of their innovations on which both models are footing. Therefore, its objective was to test if the model innovations are white, as in the case when the model completely extracts information necessary to forecast volatility.

The residual testing was explored by using Hinich triple correlation test, which is based on the fourth order cumulant function. The concept of cumulants and moments is also introduced. The results demonstrated that neither ARMA-GARCH nor RV-GARH, if based on the second order statistics, produce “white residuals.

The finding that innovations are not white has implications for modeling FX spot price dynamics. If there are both third- and fourth-order nonlinear serial dependence in the data, then time series models that make use of a linear structure, or presume a pure white noise input, such as the geometric Brownian motion (GBM) stochastic diffusion model, are problematic. In particular, the dependence structure violates both the normality and Markovian assumptions underpinning conventional GBM models.

This finding of serial dependence in innovations which can be categorized as a deep structural phenomenon, opens question one of which is, ‘Is it still appropriate to simply equate the presence of non Gaussianity and nonlinearity with the presence of outliers?’. It has important implications for the use of GBM and jump diffusion models that currently emphasize accepted risk management strategies based on the Black–Scholes option pricing model, which are employed in financincial and investment management (Hinich , 1982). Therefore, the question of parameter estimation in volatility forecasting remains an everlasting problem which definitely needs to be addressed in terms of a HOC estimation methodology.

REFERENCES

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. 2000. Exchange Rate Returns Standardized by realized volaitility are nearly Gaussian. Multinational Financial Journal. 4,159-179.

T. W. Anderson T.W. and Darling D.A. 1952. Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. Annals of Mathematical Statistics. 23(2), 169-313

Bai, N., Russell, J. R., & Tiao, G. C. (2003). Kurtosis of GARCH and stochastic volatility models with non-normality. Journal of Econometrics, 114, 349–360.

 ISBN : 9780974211428

Bollerslev, T. (1982). Generalized Autoregressive Conditional Heteroskedasticity, in *ARCH Selected Readings*, ed. by Engle, R. Onford: Onford UP, 42-60.

Box, G. E. P. and Pierce, D. A. (1970) .Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. Journal of the American Statistical Association, 65: 1509–1526. JSTOR 2284333

Corsi, F.(2009) .A Simple Approximate Long Memory Model of Realized Volatility Journal of Financial Econometrics 7: 174-196

Carrol R. and Kearvey C (2009), "GARCH Modeling of stock market volatility "in Gregoriou G. (ed), Stock Market Volatility, CRC Finance Series, CRC Presss, 71-90.

Chortareas G., Nankervis J. & Jiang Y. 2007. Forecasting Exchange Rate Volatility with High Frequency Data: Is the Euro Different? Money Macro and Finance (MMF) Research Group Conference .retrieved from : http://ideas.repec.org/p/mmf/mmfc06/79.html

Fisher R.A.1929.Moments and Product Moments of Sampling Distributions..*Proceedings of the London Mathematical Society*, Series 2, 30: 199-238 . Retrieved from : http://digital.library.adelaide.edu.au/dspace/bitstream/2440/15200/1/74pt2.pdf

Hinich, M.J. (1982), 'Testing for Gaussianity and Linearity of a Stationary Time Series', Journal of Time Series Analysis, 3: 169–76.

Hinich, M.J. (1996) Testing for dependence in the input to a linear time series model. Journal of Nonparametric Statistics 6, 205–221

Lim K.P., Hinich M. and LiewV. K. (2006): Statistical Inadequacy of GARCH Models for Asian Stock Markets: Evidence and Implications .Journal of Emerging market Finance, 4(3) Retrieved from http://hinich.webhost.utenas.edu/files/economics/asian-garch.pdf

Lobito I., Nankervis J. & Savin. N. Testing for Autocorrelation Using a Modified Box-Pierce Q Test International Economic
Review.2001.42 (1), 187-205

Poon H. and Granger C., (2003), "Forecasting Volatility in Financial market: A review", Journal of Economic Literature, Vol. XLI, 478-539

Teräsvirta T Zhao Z.(2011).Stylized Facts of Return Series, Robust Estimates, and Three Popular Models of Volatility. *Applied Financial Economics* Volume 21, Issue 1-2, pp.67-94.

Zumbach, G. O., Corsi, F. and Trapletti, A.: 2002, Efficient estimation of volatility using high frequency data, *Working Paper*