

ECONOMETRIC APPROACH OF HETEROSKEDASTICITY ON FINANCIAL TIME SERIES IN A GENERAL FRAMEWORK

FELICIA RAMONA BIRĂU

PH.D STUDENT*

UNIVERSITY OF CRAIOVA,

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

CRAIOVA, ROMANIA

e-mail: biraumona@yahoo.com

Abstract

The aim of this paper is to provide an overview of the diagnostic tests for detecting heteroskedasticity on financial time series. In financial econometrics, heteroskedasticity is generally associated with cross sectional data but can also be identified modeling time series data. The presence of heteroscedasticity in financial time series can be caused by certain specific factors, like a model misspecification, inadequate data transformation or as a result of certain outliers. Heteroskedasticity arise when the homoskedasticity assumption is violated. Testing for the presence of heteroskedasticity in financial time is performed by applying diagnostic test, such as : Breusch-Pagan LM test, White's test, Glesjer LM test, Harvey-Godfrey LM test, Park LM test and Goldfeld-Quand test.

Key words : heteroskedasticity, linear dependence, variance, volatility clustering, non-normal distribution

JEL classification : C58, G17

1. Introduction

Generally, heteroskedasticity is perceived as a specific feature of cross sectional data, but that does not mean it can not be associated with time series data. In addition, financial time series are characterized by the existence of volatility clustering, chaotic behavior and pronounced instability. Beyond these issues, financial time series data exhibit linear dependence in volatility, which implies the existence of heteroskedasticity.

In a technical manner this concept can be summarized as a violation of an Ordinary Least Squares (OLS) assumption. An essential OLS assumption is that the variance of the error terms is constant and independent or serially uncorrelated. In financial mathematics the previous assumption has the following expression :

$$\text{var}(u_t) = \sigma^2, \forall t \in 1, 2, \dots, n$$

The consequence is that OLS estimators will not be BLUE and in the light of this fact they will not be efficient, accurate or consistent. In fact, OLS estimators will be Linear and Unbiased but will not be the most precise estimators.

Practically the error term for each observation is the same for all observations. Specifically, it was assumed that having a constant variance means the disturbances are homoskedastic. At this point, it is important to emphasize the fact that heteroskedasticity and homoskedasticity are antagonistic concepts. Even the etymology of these terms of Greek origin suggests their contradiction, so homo meaning the same or equal and hetero meaning different or unequal with the same common root skedasmós meaning spread or scatter. Consequently, the mathematical formula reflecting the heteroskedasticity assumption, meaning the variance of the error terms depends on the analyzed observations or that the variance for each observation could be different, has the following expression :

$$\text{var}(u_t) = \sigma_t^2, \forall t \in 1, 2, \dots, n$$

Considering the previous arguments, testing heteroskedasticity should be performed maintaining the assumption that the error terms are actually homoskedastic. Therefore, it must be analyzed if the null hypothesis is true, respectively :

$$H_0 : \text{var}(u_t) = \sigma_t^2, \forall t \in 1, 2, \dots, n$$

In order to detected heteroskedasticity should be considered all the Ordinary Least Squares assumptions, so that specification tests on coefficient estimates could be performed properly. These assumptions are the following :

a) $E(u_t) = 0$

* ACKNOWLEDGEMENT - This work was supported by the strategic grant POSDRU/CPP107/DMI1.5/S/78421, Project ID 78421 (2010), co-financed by the European Social Fund – Investing in People, within the Sectoral Operational Programme Human Resources Development 2007 – 2013

- b) $\text{var}(u_t) = \sigma^2 < \infty$
- c) $\text{cov}(u_i, u_j) = 0$
- d) $\text{cov}(u_i, x_t) = 0$
- e) $u_t \sim N(0, \sigma^2)$

2. Diagnostic tests for detecting heteroscedasticity

In literature there are several methods for detecting heteroscedasticity. The most unpretentious, but rather inaccurate is the graphical method. Thus, it is necessary to generate a plot of residuals and independent variables. In this case, the presence of heteroscedasticity can be identified by analyzing the scatter plot. Concretely, the lack of heteroscedasticity must coincide with the absence of an obvious pattern to the spread of the disturbance term.

Results of greater accuracy are obtained by applying diagnostic tests for detecting heteroscedasticity, such as : Breusch-Pagan LM test, White's test, Glesjer LM test, Harvey-Godfrey LM test, Park LM test and Goldfeld-Quand test.. Differences between these tests are more or less significant in the context of financial time series modeling.

Breusch-Pagan LM test was developed in 1979 by Trevor Breusch and Adrian Pagan. Considering the following regression model :

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad \text{where } \text{var}(u_i) = \sigma_i^2$$

The Breusch-Pagan LM test involves a series of intermediate stages in detecting heteroscedasticity. First, it is implemented the regression of the previous equation and there are obtained the residuals $\hat{u}_i$. Subsequently, the auxiliary regression equation it is established, as the following :

$$\hat{u}_i^2 = b_1 + b_2 W_{2i} + \dots + b_m W_{mi} + v_i$$

where W_{mi} is a series of variables established to determine the variance of the error terms. The next step involves setting the null hypothesis of homoskedasticity as :

$$H_0 : b_1 = b_2 = \dots = b_m = 0$$

In the case that at least one of the b_s is different from zero and at least one of the W_s influences the variance of the error terms, the null hypothesis is rejected.

The following step is to compute the $LM=nR^2$ statistic, where n is the number of observations established to determine the auxiliary regression and R^2 is the coefficient of determination. The LM-statistic follows the χ^2 distribution characterized by $m-1$ degrees of freedom. The final step assume to reject the null hypothesis and to highlight the presence of heteroscedasticity if LM-statistical is higher than the critical value.

White's test was developed by Halbert White in 1980 and it is a generally, unrestricted and widely used diagnostic test for detecting heteroscedasticity in the residuals from a least squares regression. Practically, the White test is a test for heteroskedasticity in OLS residuals. The null hypothesis is that there is no heteroskedasticity.

First it is estimated the following equation:

$$Y_i = \beta_1 + \beta_2 \cdot X_2 + \beta_3 \cdot X_3 + u_i$$

After that it is computed the regression of the previous equation and there are obtained the residuals $\hat{u}_i$. The auxiliary regression equation it is established, as the following :

$$\hat{u}_i^2 = b_1 + b_2 G_{2i} + b_3 G_{3i} + b_4 G_{2i}^2 + b_5 G_{3i}^2 + b_6 G_{2i} G_{3i} + v_i$$

The following stage involves setting the null hypothesis of homoskedasticity as :

$$H_0 : b_1 = b_2 = \dots = b_m = 0$$

Thereby the null hypothesis H_0 highlights the fact that the variance of the residuals is homoskedastic , i.e.,

$$\text{var}(\varepsilon_i) = \text{var}(Y_i) = \sigma^2$$

The alternative hypothesis is H_1 aims the fact that the variance of the residuals is heteroskedastic.

$$\text{var}(\varepsilon_i) = \text{var}(Y_i) = \sigma_i^2$$

In the case that at least one of the b_s is different from zero the null hypothesis is rejected.

One very important step is to compute the $LM=nR^2$ statistic, where n is the number of observations established to determine the auxiliary regression and R^2 is the coefficient of determination .

The LM-statistical follows the χ^2 distribution characterized by 6-1 degrees of freedom. The last stage assume to reject the null hypothesis and to highlight the presence of heteroscedasticity when LM-statistical is higher than the critical value.

Glesjer LM test was developed by Glesjer in 1969 and it is very similar to Breusch-Pagan LM test. The only major difference relates to the auxiliary regression equation, which is :

$$|\hat{u}_i| = b_1 + b_2 W_{2i} + \dots + b_m W_{mi} + v_i$$

Except this particular issue, every single step is repeated exactly as there were presented in the Breusch-Pagan LM test. Despite this seemingly insignificant detail, after processing a financial time series, results will not be identical.

Park LM test was developed by Park in 1966. According to Gujarati, the Park LM test formalizes the graphical method by suggesting that σ^2 is a particular function of the explanatory variable X_i .

As in the case of the other diagnostic tests for detecting heteroscedasticity discussed above, Park LM test is a set of intermediate stages.

In order to obtain the error terms $\hat{u}_i$, we run a regression equation, as the following :

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad \text{where } \text{var}(u_i) = \sigma_i^2$$

The next step aims to run the auxiliary regression :

$$\ln(\hat{u}_i^2) = b_1 + b_2 \ln W_{2i} + b_3 \ln W_{3i} + \dots + b_m \ln W_{mi} + v_i$$

The null hypothesis of homoskedasticity has the following expression :

$$H_0 : b_1 = b_2 = \dots = b_m = 0$$

The alternative hypothesis is H_1 emphasizes the fact that at least one of the b_s is different from zero.

In the case that at least one of the b_s is different from zero the null hypothesis is rejected.

The following step is to compute the $LM=nR^2$ statistic, where n is the number of observations established to determine the auxiliary regression and R^2 is the coefficient of determination the regression. The LM-statistic follows the χ^2 distribution characterized by $m-1$ degrees of freedom. Finally, reject the null hypothesis and notice the presence of heteroscedasticity if LM-statistical is higher than the critical value.

Harvey-Godfrey LM test was developed by Harvey in 1976 and Godfrey in 1978. In order to performe this test, a sequence of complementary steps have to be implemented. First of all, in order to obtain the disturbances $\hat{u}_i$, we run a regression of the same initial equation like in the case of Breusch-Pagan LM test and Park LM test, respectively :

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad \text{where } \text{var}(u_i) = \sigma_i^2$$

The next step involves to compute the following auxiliary regression :

$$\ln(\hat{u}_i^2) = b_1 + b_2 W_{2i} + \dots + b_m W_{mi} + v_i$$

The null hypothesis of homoskedasticity has the following expression :

$$H_0 : b_1 = b_2 = \dots = b_m = 0$$

The alternative hypothesis is H_1 aims the fact that at least one of the b_s is not similar to zero :

In the case that at least one of the b_s is different from zero and at least one of the W_s influences the variance of the error terms, the null hypothesis is rejected.

The following step is to compute the $LM=nR^2$ statistic, where n is the number of observations established to determine the auxiliary regression and R^2 is the coefficient of determination of this particular regression. The LM-statistical follows the χ^2 distribution characterized by $m-1$ degrees of freedom. The last step assume to reject the null hypothesis and to highlight the presence of heteroscedasticity when LM-statistical is higher than the critical value.

Goldfeld-Quand test was developed in 1965 by Stephen Goldfeld and Richard Quandt and is a solid alternative to LM tests highlighted above. Applying this test requires to performe a sequence of intermediate stages. First step involves to arrange the observations either in ascending or in descending order. For example, let us consider that the option is to arrange the data in ascending order, from the lowest to the highest value of the independent variable X_i that is supposed to generate heteroscedasticity. The next stage aims to divide the ordered sequence into two equal sub-sequence by omitting an arbitrary number p of the central observations. Consequently, the two equal sub-sequence

will summarize each of them a number of $\frac{1}{2}(n-p)$ observations. Therefore, compute two different OLS regressions,

the first one for the lowest values of X_i and the second for the highest values of X_i . In addition, obtain the RSS for each regression equation, RSS_1 for the lowest values of X_i and RSS_2 for the highest values of X_i . An F-statistic is calculated based on the following formula :

$$F = \frac{RSS_1}{RSS_2}$$

The F-statistic is distributed with $\frac{(N-p-2k)}{2}$ degrees of freedom for both numerator and denominator.

Subsequently compare the value obtained for the F-statistic with the tabulated value of F-critical for the specified number of degrees of freedom and a certain confidence level. If F-statistic is higher than F-critical, the null hypothesis of homoskedasticity is rejected and the presence of heteroskedasticity is confirmed.

A significant role in the proper use of Goldfeld-Quand test lies on the choice of the arbitrary number p which should generally be based on the rule of thumb. Accordingly, it is practiced to eliminate from the analysis between 1/6 and 1/3 of the observations.

Although Goldfeld-Quandt test is uncomplicated and widely used, especially for simple regressions, it has several disadvantages, such as : is a quite intuitive parametric test, provides a very low accuracy in the case of

unknown or unobserved explanatory variables and it is characterized by a slight robustness regarding specification error terms.

3. Conclusions

Financial time series data exhibit an atypical and chaotic behavior characterized by instability, uncertainty, subjectivity, volatility clustering especially in the context of a global financial crisis. In addition, financial time series data exhibit linear dependence in volatility, which implies the existence of heteroskedasticity. Concretely, violating the homoskedasticity assumption determines implicitly the presence of heteroskedasticity.

Heteroskedasticity can arise as a consequence of various circumstances and there are several methods for detecting its presence. In this respect, there are mentioned the graphical method and the diagnostic tests for detecting heteroscedasticity, respectively : Breusch-Pagan LM test, White's test, Glesjer LM test, Harvey-Godfrey LM test, Park LM test and Goldfeld-Quand test.

Practically, testing for the presence of heteroskedasticity in financial time series do not involve special difficulties, as long as it takes into account the characteristics of the analyzed time series data. Despite the fact that heteroskedasticity is generally associated with cross sectional data, there are circumstances when it becomes a feature of time series data. Heteroscedasticity can arise in financial time series due to the influence of certain factors such as : a model misspecification, inadequate data transformation or as a result of certain outliers.

Detecting heteroskedasticity is an important issue that must be considered in the context of financial time series modeling and forecasting, especially considering the interests of potential investors.

4. Bibliography

- [1] **Bollerslev, T., Engle, R.F.**, Common persistence in conditional variances, *Econometrica*, Vol. 61, No. 1, 1993, 167-186
- [2] **Dickey, D.A.**, Stationarity Issues in Time Series Models, *Statistics and Data Analysis*, Philadelphia, 2005, paper 192-30.
- [3] **Dwyer, G.P.**, Nonlinear Time Series and Financial Applications, Federal Reserve Bank of Atlanta, January 2003
- [4] **Gujarati, D.N.**, Basic Econometrics, Fourth Edition, Published by McGraw-Hill/Irwin, 2003, ISBN: 978-0-07-233542-2
- [5] **Lineesh, M.C. et al.**, Analysis of Nonstationary Nonlinear Economic Time Series of Gold Price. A Comparative Study, *International Mathematical Forum*, 5, India, 2010, no. 34, pp.1673 - 1683
- [6] **Metzner, P. et al.**, Analysis of persistent non-stationary time series and applications, *Institute of Computational Science*, University of Lugano, 2012
- [7] **Mishra, PK et al.**, Stationarity of Stock Return Series: The Case of India, *European Journal of Scientific Research*, EuroJournals Publishing, Inc., Vol.44 No.4, 2010, pp.596-601, ISSN 1450-216X
- [8] **Nachane, D.**, Selected Problems in the Analysis of Non-stationary & Nonlinear Time Series Quantitative Approaches to Public Policy, *IGIDR Proceedings/Project 2009*, <http://www.igidr.ac.in/pdf/publication/PP-062-01.pdf>
- [9] **Rodriguez, J., Ruiz, E.**, A powerful test for conditional heteroscedasticity for financial time series with highly persistent volatilities, *Statistica Sinica* 15, 2005, pp.505-525
- [10] **Scott Long, J., Ervin, L.**, Correcting for Heteroscedasticity with Heteroscedasticity Consistent Standard Errors in the Linear Regression Model: Small Sample Considerations, *Indiana University Bloomington*, IN 47405, 1998
- [11] **Strozzi, F.**, Application of non-linear time series analysis techniques to high frequency currency exchange data, *Liuc Papers n. 99*, *Serie Metodi quantitativi* 14, Suppl. a gennaio 2002
- [12] **Thomson, D.J.**, Jackknifing multiple-window spectra, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, VI, 73-76, 1994
- [13] **Yüce, M.**, An asymptotic test for the detection of heteroskedasticity, *Ekonometri ve İstatistik Sayı:8*, 2008, pp. 33-44
- [14] **Wickremasinghe, G.B.**, Efficiency of foreign exchange markets: A developing country perspective, *Asian Academy of Management Journal*, Vol. 10, No. 1, 1-17, January 2005