Stock Index Forecasting for Vietnam's Stock Market

Son Van Duong[1], Jui Fang Chang[2]

[1]International Business Department of National Kaohsiung University of Applied Sciences

[2]Manufacturing and Management Department of National Kaohsiung University of Applied Sciences

Kaohsiung 807, Taiwan

hisa_son@yahoo.com.vn, rose@cc.kuas.edu.tw

Abstract *In the so-called global crisis from late 2007 up to early 2009, the economy was shaken, especially stock market by global increased volatility transmission. Stocks' price in Viet Nam continuously went down. Therefore, in this paper, we will analyze to find main factors which strongly have influence on fluctuation of Vietnam's stock index for the period from 2009 to 2011. Then, we estimate the regression model, GARCH, GARCH in mean, EGARCH, EGARCH in Mean and GJR-GARCH and put them into the comparison to find the best fit model to forecast Vietnam's stock index. Using the mean square error (MSE), Mean absolute error (MAE) and Root mean square error (RMSE) criterion to evaluate and compare models, we choose Regression-EGRCH(1,1,1)-M which has the superior forecast ability.*

Keyword: Regression, GARCH, GARCH-M, EGARCH, EGARCH-M, GJR-GARCH model; MSE, RMSE and MAE criteria.

I. INTRODUCTION

After the crisis in the period of late 2007 to early 2009, Vietnam's stock market is significantly going down. How to help it to recover and stably develop is a big issue now. One of solution is to predict the trend of stock exchange then make the appropriate decision to control the market effectively. Due to the increasing of globalization, international investments and the worldwide circulation of capital results in close relationships between stock markets, we analyze the correlations between Vietnam's stock index (represented by VNINDEX) and other indices S&P500 (US), CAC (France), DAX (Germany), FTSE100 (UK), KOSPI (Korea), STRAITS TIMES (Singapore) based on Spearman's correlation coefficient theory. With the significant correlations between VNINDEX and other six indices (Table 1 and Figure 1), we go to find the most suitable model to forecast VNINDEX based on other six indices above. Regression model allows to regression one dependent variable on other independent valuables. Then linear models, ARIMA (Box-Jenkins, 1976), ARCH (Angle, 1982), and GARCH (Angle, 1995), were introduced with better forecast ability. Nonlinear GARCH (the extensions of GARCH) continued having been proposed which are Exponential GARCH (EGARCH) (Nelson, 1991), GJR–GARCH (Glosten etal., 1993) to capture the leverage effect that positive and negative shocks have an *asymmetric* impact on the conditional volatility of subsequent observations. This paper combines regression models with GARCH, EGARCH and GJR-GARCH and their alternatives to generate the best fit model for forecasting VNINDEX.

Table 1: Correlations between VNINDEX and other six indices for the period 2009-2011

Correlations	
	VNIDEX
CAC	0.69398
DAX	0.47935
S&P500	0.44624
FTSE100	0.53879
KOSPI	0.41770
STRAITS TIMES	0.65346

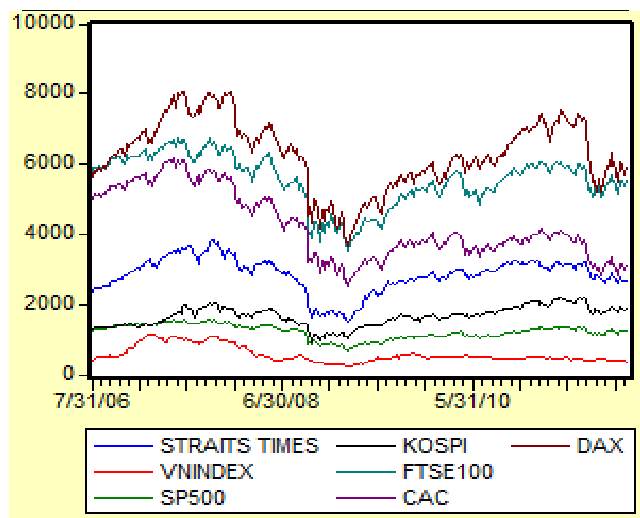

Figure 1: The Graph of six stock price indices VNINDEX, CAC, DAX, STRAITS TIMES, KOSPI, FTSE100, S&P500 from 2006 to 2011

II. METHODOLOGY

In the field of modeling time series in finance, beside the classical model, regression, there are some similar models introduced like ARIMA, ARCH, GARCH and they were enriched. It leads to model choice has become more complicated. In this paper, we suggest using regression model and combining regression with GARCH, EGARCH and GJR-GARCH. And the alternatives are GARCH-M and EGARCH-M.

1. Regression model

$Y_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon_t$ (1)

Where:

- The unknown parameter denoted as β
- The independent variables, X
- The dependent variables Y_t.
- ε_t is an error term and the subscript *t* indices a particular observation. ε_t is the difference between the actual observed value Y_t and the average value or mean $E(Y/X_t)$.

2. Regression-GARCH model

$$Y_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon_t \quad (2)$$

With $\varepsilon_t \sim N(0, h_t)$.

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}^2 \quad (3)$$

Where h_t is the variance of ε_t

The variance h_t in the GARCH equation (2.3) not only depends on the past value of socks but also the past value of variance itself.

3. Regression-GARCH –M model

$$Y_t = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + \delta h_t + \varepsilon_t \quad (4)$$

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}^2 \quad (5)$$

With t,j = 1,…,n

The variance, h_t of ε_t now is added into the formula (4) and It's called GARCH in mean. It means that a stock market with higher risk should have a higher increase in stock index.

4. Regression-EGARCH model

$$Y_t = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + \varepsilon_t \quad (6)$$

$$\mathrm{Ln}(h_t) = \alpha_0 + \alpha_1(\varepsilon_{t-1}/h_{t-1}^{0.5}) + \lambda_1|\varepsilon_{t-1}/h_{t-1}^{0.5}| + \beta_1 \ln(h_{t-1}) \quad (7)$$

The model allows for the asymmetric effect of news is the exponential-GARCH (EGARCH) model. Nelson (1991) proposed this kind of specification to eliminate the requirement of non-negativity constraints of GARCH.

5. Regression-EGARCH-M model

$$Y_t = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + \ln(h_t) + \varepsilon_t \quad (8)$$

$$\mathrm{Ln}(h_t) = \alpha_0 + \alpha_1(\varepsilon_{t-1}/h_{t-1}^{0.5}) + \lambda_1|\varepsilon_{t-1}/h_{t-1}^{0.5}| + \beta_1 \ln(h_{t-1}) \quad (9)$$

EGARCH in Mean model has the same idea with GARCH in mean but the variance of the residual added into the expected value equation is in the form of natural logarithm.

6. Regression-GJR-GARCH model

$$Y_t = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + \varepsilon_t \quad (10)$$

$$h_t = \alpha_0 + \sum_{j=1}^{q} (\alpha_j + \gamma s_{t-j}) \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j},\ t = 1,\ldots,n \quad (11)$$

Where: s_{t-j} is a dummy variable. $s_{t-j} = 1$ if $\varepsilon_{t-j} < 0$ and $s_{t-j} = 0$ if $\varepsilon_{t-j} \geq 0$.

The main purpose of the model is to consider the nature of asymmetric shocks between negative and positive shocks. If the coefficient γ is significant, negative and positive shall have different impacts on the variance. In detail, a bad news merely affects α_j, while a negative news can affect $\alpha_j + \gamma$.

In each model above, parameter β_i represents the impact of an independent variable at time t towards dependent valuable at time t. β varies over specific models because of different estimation method of residual in each type of model with it respective characteristic.

7. Evaluation of model

We will evaluate and compare the forecast results by using three criteria:

-Mean square error: $$\mathrm{MSE} = \sum \frac{(\hat{S}_t - S_t)^2}{n} \quad (12)$$

-Mean absolute error: $$\mathrm{MAE} = \sum \frac{|\hat{S}_t - S_t|}{n} \quad (13)$$

-Root Mean square error: $$\mathrm{RMSE} = \sqrt{\sum \frac{(\hat{S}_t - S_t)^2}{n}} \quad (14)$$

III. EXPERIMENT RESULTS

1. Appropriate models

Data analysis, hypotheses test and criteria evaluation was conducted. Three appropriate models were generated as follows:

1.1 Regression model:

All probability of regressive coefficients are <5%, adjusted R-squared = 0.81, the residual has a standard normal distribution and MSE, RMSE, MAE are within 0 and 1(Table 2) indicate the goodness of fit of the regression model.

Table 2: MSE, RMSE, MAE values of residual in regression model

	Value
RMSE	**0.08025**
MSE	**0.00644**
MAE	**0.06365**

The regression model is constructed below:

$Log(VNINDEX_t) = 1.202log(S\&P500) - 0.726log(KOSPI) + 1.184log(CAC) - 1.114log(DAX) - 1.081log(FTSE100) + 1.558log(STRAITS\ TIMES)$ (15)

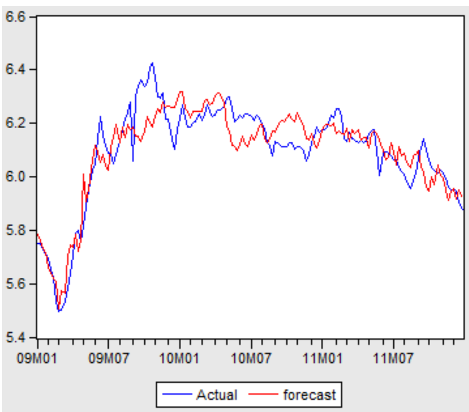

Figure 2: The graph of actual and regression model- forecasted results

1.2 Regression GARCH(1,1)-M:

All probability of regressive coefficients and in variance equation are <5%, adjusted R-squared = 0.87 and MSE, RMSE, MAE are within 0 and 1(Table 3) indicate the goodness of fit of the model.

Table 3: MSE, RMSE, MAE values of residual in GARCH(1,1)-M model

	Value
RMSE	0.12735
MSE	0.01622
MAE	0.08445

The regression GARCH(1,1)-M model is constructed below:

$Log(VNINDEX_t) = 1.446log(S\&P500) - 0.707log(KOSPI) + 1.110log(CAC) - 1.031log(DAX) - 1.987log(FTSE100) + 1.273log(STRAITS\_TIMES) + 0.098log(h_t)$ (16)

And the variance h_t was estimated below:

$h_t = 0.000756 + 0.271\varepsilon_{t-1}^2 + 0.568h_{t-1}$ (17)

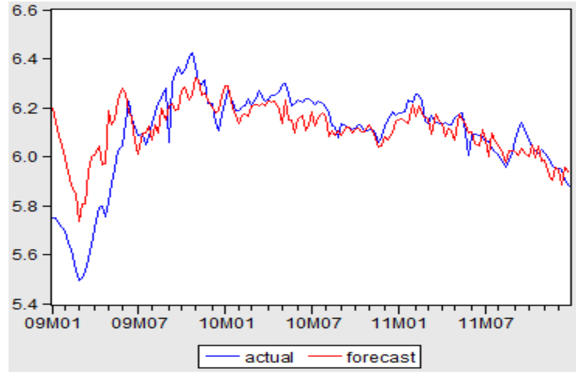

Figure 3: The graph of actual and GARCH(1,1)-M model -forecasted results

1.3 Regression EGARCH(1,1)-M:

All probability of regressive coefficients and in variance equation are <5%, adjusted R-squared = 0.91 and MSE, RMSE, MAE are within 0 and 1(Table 4) indicate the goodness of fit of the model.

Table 4: MSE, RMSE, MAE values of residual in EGARCH(1,1,1)-M model

	Value
RMSE	**0.07038**
MSE	**0.00495**
MAE	**0.04579**

The regression EGARCH(1,1,1)-M model is constructed below:

$Log(VNINDEX_t) = 1.023log(S\&P500) - 0.596log(KOSPI) + 0.848log(CAC) - 1.207log(DAX) - 0.651log(FTSE100) + 1.457log(STRAITS\_TIMES) + 0.100log(h_t)$ (18)

And the variance h_t was estimated below:

$Log(h_t) = -1.128 + 0.304(\varepsilon_{t-1}/h_{t-1}^{0.5}) + 0.215|\varepsilon_{t-1}/h_{t-1}^{0.5}| + 0.841log(h_{t-1})$ (19)

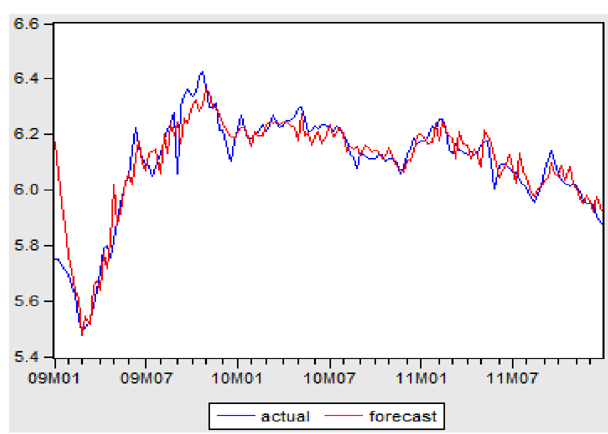

Figure 4: The graph of actual and EGARCH(1,1,1)-M model -forecasted results

2. Model comparison

With four models used above to forecast VNINDEX based on other six indices: Regression model; GARCH; EGARCH; GJR-GARCH, only GJR-GARCH is not appropriate to be applied. Other three models are fit for forecasting. However, to decide which one is the best fit model in this situation, now the researcher continues doing comparison among these models.

Except for the measurement of MSE, RMSE and MAE, the results show that EGARCH(1,1,1)-M has superior ability. We also found that all the difference in MSE, RMSE and MAE between regression model and EGARCH(1,1,1)-M model are very small.

With 156 observations, the criteria and graph in comparison of three models above are below:

Table 5 Comparison of experiment results 2

Observations	MSE	RMSE	MAE
Regression model	0.00644	0.08025	0.06365
GARCH(1,1)-M	0.01622	0.12735	0.08445
EGARCH(1,1,1)-M	***0.00495***	***0.07038***	***0.04579***

By glancing at graphs above, it is difficult to see which model has superior forecast performance, because all three models have small difference between actual and forecasted values in terms of residual. It means all three models above have similar forecast ability. But if a careful look is taken, we easily find that EGARCH(1,1,1)-M model provides the best forecasted value because the residual is smallest.

IV. CONCLUSION AND SUGGESTION

Our conclusion comprises two components. First, when doing forecast based on regression theory, we should analysis the actual economy to determine main factors which strongly have influence on VNIDEX. Second, using many models then testing and evaluating them to generate the best fit model. Among models above, EGARCH-M with the valuable innovation is the best fit model.

Future researchers could using more other models such as Genetic Algorithms (GA), genetic algorithm, back propagation network (BPN) and back propagation network (GABPN), etc. and put all of them in comparison to generate the best fit model.

REFERENCE

[1] Tsay, R. S. (2004). *Analysis of Financial Time Series*. New York, USA: John Wiley & Sons, Inc.

[2] Nelson, D. B. (1990). Stationarity and persistence in the GARCH(1, 1) model. Econometric Theory, 6, 318-334.

[3] Amemiya, T. (1985): *Advanced Econometrics*. Cambridge: Harvard University Press.

[4] Basrak, B., Davis R.A., and Mikosch, T. (2002): Regular variation of GARCH processes. *Stochastic. Process. Appl*., 99, 95-115.

[5] George, L. (2004). Computational finance: Numerical methods for pricing financial instrument. Oxford: Oxford University.

[6] Yahoo Corporation. (2011). Get Quotes. Available at http://finance. yahoo. Com /q/hp?s=^ FTSE+Historical+Prices [Accessed 10 Jan. 2012].

[7] Wall Street Security. (2011). Stock code search. Available at http://wss.com. vn/tabid/292/ default.aspx? Search=Stock Code%3D0|StartDate%3D01/01/1900|EndDate%3 D-1 [Accessed 10 Jan, 2012].

[8] Kugler, P., and Lenz, C., "Chaos, ARCH and The Foreign Exchange Market: Empirical Results From Weekly Data," Unpublished Manuscript, Volkswirtschaftliches Institut, Zurich, 1990.

[9] Engle, R. and V. Ng (1993). Measuring and Testing of the Impact of News on Volatility. Journal of Finance, 48, 1749-1778.

[10] Bollerslev, T., Chou, R., and Kroner, K., "Arch Modeling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, April-May 1992, pp. 5-59.