Analysis of China Stock Market: Volatility and Influencing Factors

Song Xiao-mei
Economics and Management School
Beijing Forest University
Beijing, China

Pan Huan-xue
Economics and Management School
Beijing Forest University
Beijing, China

Abstract—**The volatility of stock market return is the main technique measurement in the risk management. This paper chooses GARCH class models to estimate in-sample period and to forecast out of sample period. The return of daily data was collected from Shanghai Stock Exchange, during the time period from January 1st 1997 to April 30th 2007. Three error measurement methods: ME, MAE and RMSE were used to evaluate the forecasting ability of GARCH class models. The empirical result indicates that EGARCH-M model is the best one for estimation of in-sample period. Asymmetric model or simple GARCH model is explained better for out of sample forecasting.**

Keywords- GARCH class models, Volatility, Forecasting

I. INTRODUCTION

The China stock market is a new and high-growth market, and is becoming influential in global economy. Furthermore, China will introduce their first stock index futures soon, so that better understanding of the volatility of China stock price is very important for the pricing derivative securities and for the measuring stock market risk. There are already existing papers about volatility in China stock market by using data in different periods, but few papers have disclosed the influencing factors that may affect the China stock volatility. This present paper will not only evaluate the performance based on some models in terms of forecasting China stock volatility but also disclose valuable data of influencing factors to analyze the reasons for China stock market volatility.

II. LITERATURE REVIEW

Many previous papers supported that ARCH models is good for forecasting stock return volatility both in China and other countries' market. The first major research of China stock market was made by Yu (1994), he described an ARCH (2) model for the Shenzhen index, and a GARCH (1, 1) model for the Shanghai index return. Liu, Song and Romily (1997) used data from 1992 to 1996 and found that both Shanghai and Shenzhen return series can be explained by the GARCH-M (1,1) model. Copeland and Zhang (2003) took over the data from 1994 to 2001 of Shanghai and Shenzhen daily index data, and concluded that Shenzhen stock market has a higher volatility than that in Shanghai market by using EGARCH models. Applying foreign countries' data by ARCH/GARCH models, Koopman and Uspensky (2002) showed that GARCH-M model proved a better fit of the volatility series for Japan stock market data from 1988 to 1998. Booth and GREGORY (1998) found that major European stock market follows EGARCH model.

Although many studies indicated that GARCH models seem to be the most successful models for forecasting the variance of return, it may be based for skewed time series. Wei (2002) suggested to use two non-linear GARCH models: one is Quadratic GARCH model (QGACH) proposed by Engle and Ng（1993,and the other is Glosten, Jagannathan and Runkle models (GJR) proposed by Glosten(1992). Alternatively, the EGARCH model can generate skewed time series, but it is not useful for repeated forecasting exercises and the parameter estimation could be tedious. Wei (2002) compared the forecasting performance of linear GARCH models with QGACH and GJR models, and found QGACH model can improve linear GARCH model, but GJT model is not recommended. Therefore,as long as more newly developed econometric techniques emerge, people may be able to obtain better forecasting models for volatility analysis.

The influencing factors may affect stock volatility, those factors including macroeconomics, government policy, financial statements and even the behavior of stockholders. Some studies show that interest rate and CPI are correlated with stock price volatility. Mueller (2006) pointed out that increased interest rate should affect stock price decrease. Dinenis and Staikouras (1998) applied UK stock return data from 1989 to 1995 and found the sensitivity of stock returns to the announcement of interest rate changes. Liljeblom and Stenius (1997) pointed out that CPI is negatively related to the volatility of stock Market in Finland. The similar results were also founded in US stock market by Valckx (2004). It shows that changes for inflation, a leading indicator, entail strong revisions in the equity premium. Kim and Nofsinger (2002) studied the behavior of individual investors in Japan from1975 to 1997 and found that buying behaviors of past winners is stronger during the bull market, but it is the opposite during the bear market. Chen, Kim, Nofsinger and Rui (1997) compared Chinese investors with US investors and found that Chinese shareholders make some poor trading decisions. Those poor decisions may cause an unusual volatility of stock market.

III. METHODOLOGIES

Since ARCH models are doing a good job for describing the characters of financial time series volatility, so that the

978-1-4244-5326-9/10/$26.00

main test technique, ARCH with following models: GARCH, TARCH, EGARCH and GARCH-M, have been used in this paper.

A. GARCH model

Based on the Akaike's information criterion (AIC) criterion, all GARCH class models have p=1 and q =1. The GARCH (1, 1) model can be written as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (1)$$

The conditional volatility today depends on yesterday's conditional volatility and squared error. The sufficient restrictions ensure that the condition variance of the return is always positive. The value of the parameters $\alpha_1 + \beta_1 < 1$ defend the condition of stationary. However, the good news and bad news may have different effects on the stock volatility.

B. TARCH model

The TARCH (1, 1) model is written as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \gamma \varepsilon_{t-1}^2 I_{t-1}. \quad (2)$$

where $I_{t-1} = 1$, if $\varepsilon_{t-1} < 0$, $I_{t-1} = 0$, if $\varepsilon_{t-1} > 0$.

This accepts different impact of good news ($\varepsilon_{t-1} > 0$) and the bad news ($\varepsilon_{t-1} < 0$) on conditional volatility according to the sign of error terms.

C. EGARCH model

The EGARCH (1, 1) model is represented as follows:

$$\log \sigma_t^2 = \alpha + \beta \log \sigma_{t-1}^2 + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \delta \left[\frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} - \sqrt{\frac{2}{\Pi}} \right] \quad (3)$$

Clearly, in the EGARCH model, the conditional variance depends on both the sign and the size of lagged residuals. In the model, γ captures the asymmetric effect. However, the Chinese companies may be less sensitive to fluctuations in their leverage ratio, since many Chinese companies are still partially owned by central or local governments.

D. GARCH-M model

The GARCH-M allows the conditional variance into the conditional mean equation, so that the expected returns is related to the expected risk. It is a time-varying risk premium model. It could be written as:

$$\sigma_t^2 = \alpha_0 + \alpha(L) \varepsilon_{t-1}^2 + \beta(L) \varepsilon_{t-1}^2 \quad (4)$$

E. Model forecast evaluations

To compare the forecasting performance of GARCH class models, previous papers have used a variety of statistics methods. This paper will use following measurement methods: mean error (ME), mean absolute error (MAE) and the root mean squared error (RMSE).

$$\text{ME} = \frac{1}{\text{k}} \sum_{t=1}^{k} (\hat{\sigma}_t^2 - \sigma_t^2) \qquad \text{MAE} = 1/k \sum_{t=1}^{k} \left| \hat{\grave{\text{O}}}_t^2 - \grave{\text{O}}_t^2 \right|$$

$$\text{RMSE} = \sqrt{1/k \sum_{t=1}^{k} (\hat{\grave{\text{O}}}_t^2 - \grave{\text{O}}_t^2)^2} \quad (5)$$

Where k is the number of observations, σ and σ denote the forecast volatility and realize volatility respectively.

To compare the performance of various models, we need to compare forecast volatilities with actual volatilities. To find out the true volatility of daily returns, this paper uses a common method. For conditional variance is defined as: $\sigma^2_t = (r_t - r')^2$ where t is from out of sample period, r_t is the real observed daily return and r'is the mean of daily return. For a related high frequency observations, the r'is close to zero, so the conditional variance can be estimated by squared daily return : $\sigma^2_t = r_t^2$. Since by using high frequency data can make this approach more close to actual value, people may try to use hourly returns. This paper will use the daily returns since it is the most frequency available data we can find. The daily returns were calculated as the change in logarithm of price: Rt = Ln(Pt) – Ln (Pt-1).

IV. DATA ANALYSIS

This paper analyzes the daily Shanghai Stock Composite Index (SHEC) from the DataStream database over a sample period from January 1st , 1997 to April 30th, 2007, totally 2560 observations. The reason for choosing the time period after 1996 is because a stock trading policy was changed on December 14th 1996, the growth and drop rate for one stock is limited 10% on one day period. The first time frame in-sample period is from January 1st , 1997 to June 30th, 2004 (1873 daily returns), and out of sample data are from July 1st ,2004 up to June 30th , 2005 (243 daily returns). Based on an increased data for one more year, doing the second time, in-sample period is chosen from January 1st, 1997 to June 30th, 2005 (2116 daily returns), and out of sample data is from July 1st, 2005 until April 30th, 2007 (444 daily returns). The reason for doing estimation and forecasting twice is because starting in Jun 2005, the Chinese government decided to allow companies giving their non-trading shares to the trading shareholders. The result is that after July, 2005, China stock market enters into a high growth bull .Therefore, this paper tries to compare the GARCH class models forecasting ability before and after high growth happened.

V. EMPIRICAL RESULTS

A. In-sample estimation of GARCH type model from 1997 to 2004

The GARCH class volatility is performed with Eviews 5.0. For estimation of GARCH class models, we start with a general specification of the mean equation and variance equation from sample period January 1st, 1997 to June 30th, 2004. If the model is correct then the standardized residuals should be independent and identically distributed with mean zero and variance one. After caculate, Q statistics of

GARCH(1,1) satisfied the null hypothesis that the standardized residuals are independent. After doing Histogram-Normality Tests, the results indicate that the standardized residuals are non-normally distributed, so we assumed a student t-distribution to approve satisfaction of misspecification test of GARCH model.

The GARCH (1,1) model is generally used to construct multi-period forecasts of volatility, so we start with this model. Then we examine the effect of risk on returns by using other GARCH class models, such as TARCH (1,1) EGARCH(1,1), EGARCH-M(1,1) and TARCH-M(1,1). First, we provide sample data from January 1st to June 30th 2005. The Akaike info criterion (AIC) and Schwarz Criterion (SC) for each of the five forecasting models is reported to compare them. Furthermore, the p values are reported in parentheses in order to check the empirical validity of above five models. The GARCH(1,1) coefficients, 0.217438 and 0.73641 are statistically significant. The sum of these coefficients is 0.95385, close to unit root. This indicates that shocks to volatility have a persistent effect on the conditional variance. All the parameters in GARCH(1,1) model are significant at 5% level, so the constant variance models are rejected within in-sample periods, and GARCH(1,1) model seems to do a better performance. From the Sign Bias Test that was proposed by Engle and Ng (1993), the p-value indicates that the null hypothesis of no sign and size bias is strongly rejected. Therefore, the volatility depends on both the sign and size of past shocks, and there exist asymmetries in volatility. For TARCH(1,1) model, the coefficient on the asymmetry term is positive (0.212017) and statistically significant at 5% level. This indicates that there may present some leverage effect for the daily return series of stock index. To apply EGARCH(1,1) model, the coefficient on the asymmetry term, -0.094163, is negative and statistically significant at 5% level. Again, this significant parameter estimate implies that leverage effect exist for the index return series. We then examine the effect of risk on returns using EGARCH-M(1,1) and TARCH-M(1,1) model. By comparison of AIC and SC, it displays that EGARCH-M(1,1) model has the smallest both AIC and SC. Therefore, the EGARCH-M (1, 1) model seems better than other GARCH class models. This result is consistent with the earlier results of significant leverage effect in daily return series and also consistent with the report of Liu, Song and Romily (1997), which indicated that Shanghai return series can be explained by GARCH-M (1,1) model.

B. In-sample estimation of GARCH type model from 1997 to 2005

Redoing the GARCH class estimation by using the sample data period from January 1st 1997 until April 30th 2005, with student t-distribution, we get the result. The result shows that the standardized residuals are still independent, so that the GARCH class models fit data very well. And it also displays that the sum of GARCH coefficients is 0.95003 now, which is slight smaller than previous in-sample result 0.95385 and is close to unity. Thus, the shocks to volatility have a persistent effect on the conditional variance. From TGARH(1,1) and EGARCH(1,1) models, we can see that the coefficient on the asymmetry term is positive and negative, respectively, which means there exist some leverage effect for the daily return of stock index. This result is consistent with previous in-sample result as well. According to AIC and SC value, the EGARCH-M (1,1) model is favorite due to the comparison of other GARCH type models. It should be recognized that the AIC and SC value is a little bigger than our previous in-sample estimation. Over all, the in-sample period is one year more now, but the conclusion of EGARCH-M(1,1) model is still the best one without change.

C. Out of sample forecasting results from 2004 to 2005

Above five estimated models are then used for forecasting the values of the index return for the one year sample period which is from July 1st 2004 to June 30th ,2005. The three measurement methods of forecast errors – ME, MAE and RMSE are computed for each forecast. Table 1 shows that the result of three measurement methods for these five models. The table also presents the ranking of the 5 forecasting models from the best to the worst. Based on the results, all GARCH class models over predict volatility. From the rankings of the models in table 1, TARCH-M(1,1) is both the best one in terms of ME and MAE, whereas, EGARCH(1,1) model is the best based on RMSE. Since all EGARCH(1,1) and TARCH-M(1,1) models are asymmetric GARCH models, they present some leverage effect, which is consistent with our best in-sample model EGARCH-M(1,1). The EGARCH-M(1,1) model is the second best model based on RMSE and the TARCH ranks the second both from ME and MAE. The GARCH(1,1) model is generally the worst one based on both MAE and RMSE. However, the ME test statistics indicates that EGARCH (1, 1) is the worst model. Again, ME test may lead to an unreliable ranking, so that we just briefly discussed this test. This result is consistent with many other stock markets. See for example Yu (2002) for the New Zealand market. Although it is hard to tell either TARCH-M model or EGARCH model is superior for forecasting performance, asymmetric GARCH type models are better than simple GARCH model.

TABLE I. OUT OF SAMPLE GARCH CLASS MODELS RESULT FROM 2004 TO 2005

SHEC	ME	MAE	RMSE
GARCH	3.11614E-05 [3]	0.000252081 [5]	0.000517451 [5]
TARCH	2.11034E-05 [2]	0.000244094 [2]	0.000505269 [3]
EGRCH	3.22489E-05 [5]	0.000244981 [4]	**0.000500738 [1]**
EGARCH-M	3.13462E-05 [4]	0.000244858 [3]	0.000501304 [2]
TARCH-M	**1.9768E-05 [1]**	**0.000243534 [1]**	0.000505956 [4]

Note: The number in [] indicate ranking of the models in each error measurement method

The actual volatility and forecasting conditional volatility from GARCH class models shows that the actual volatility in out of sample period is increasing as time going on. The range of volatility is from 0 to 0.007. Most representatively, in September 2004, January 2005 and May 2005, the volatility had a big jump. Among GARCH class models, the best graphs are TARCH (1,1) and TGRCH-M (1,1) model, since both conditional volatility ranges are from 0 to 0.012, which is close to the actual volatility graph. Furthermore, the TARCH (1,1) and TGRCH-M (1,1) graphs are all above that in the three time periods with big jumps as described. It's hard to tell which model is the best choice based on the graphs.

Therefore, based on the graph analysis and above forecast errors measurement, we may prefer TARCH-M (1,1) to be our forecasting model for the return volatility.

D. *Out of sample forecasting results from 2005 to 2007*

Starting July 1st, 2005, China stock index has increased rapidly, so we want to see how the GARCH class models performed during this big bull period. In this paper we analyzed the forecasting of the index return for the period of July 1st 2005 to April 30th 2007. Table 2 shows the statistic results carried out on forecasting by different GARCH type models. All GARCH models show over predication of volatility except for EGARCH-M (1,1) model that gives an under prediction. Furthermore, EGARCH-M (1,1) is the model based on ME test. According to more reliable MAE and RMSE test statistics, the GARCH (1,1) model is ranking number one. In fact, the RMSE ranking is the mostly reinforced results obtained from MAE, both RMSE and MAE indicates that TARCH (1,1) model is the worst one. In addition, all MAE and RMSE results are bigger than the previous out of sample results during the period from July 1st, 2004 to June 30th, 2005. The reason for this may be due to the high volatility growth starting July 2005, so the forecasting ability is not good as previous out of sample period. It seems that GARCH (1,1) model is better for forecasting ability than other asymmetric models, and it is also consistent with the model described by Yu (1994). Again, it is difficult to choose one single super model from different kinds of error statistics; but the results show that the leverage effect may not exist during the big bull period. Alternatively, the out of sample period is ten months longer than the previous out of sample period. This may reduce the GARCH class model's forecasting ability, so that the GARCH (1,1) model is still dubitable.

TABLE 2 OUT OF SAMPLE GARCH CLASS MODELS RESULT FROM 2005 TO 2007

SHEC	ME	MAE	RMSE
GARCH	3.58213E-06 [3]	0.000263629 [1]	0.000555385 [1]
TARCH	1.58129E-05 [5]	0.000276991 [5]	0.00057878 [5]
EGRCH	1.47328E-06 [2]	0.000265258 [3]	0.000560072 [2]
EGARCH-M	-9.80012E-07 [1]	0.000264101 [2]	0.000560409 [3]
TARCH-M	1.35454E-05 [4]	0.000275713 [4]	0.000577665 [4]

Note: The number in [] indicate ranking of the models in each error measurement method

The actual volatility and forecasting conditional volatility from GARCH class models shows that Clearly both actual volatility and GARCH class models' volatility are increased. The actually volatility range is from 0 to 0.009 and bigger than the previous out of sample test from July 1st 2004 to June 30th 2005. The most volatility period happened in July 2006 and February 2007. Comparing the GARCH class model forecasting result, volatility range by all models becomes smaller than the actual range, which means during the high rapidly growth period, the forecasting model may not perform well.

E. *Empirical results summary*

Since compared with volatility graphs are not a technique way to consider which model is fitted data very well, so we compared AIC and SC value by different models. The result indicate that EGARCH-M (1,1) is a better estimation model for in-sample time period. For out of sample, we used ME MAE and RMSE error statistics to define forecasting ability. Based on out results and above analysis, a asymmetric GARCH type models, namely EGARCH-M (1,1) and TARCH-M(1,1) are better than the simple GARCH model for the data collected from 2004 to 2005 forecasting period. The GARCH (1,1) model may fit better for the data during 2005 to 2007 period.

VI. THE INFLUENCING FACTORS FOR VOLATILITY

From above analysis, we can see that China stock market has a high volatility, but what factors may affect stock volatility and what is China market's unique characteristic? This paper demonstrates the relationship between the stock market volatility and some macroeconomic factors by analyzing interest rate, CPI and government policy.

A. *Interest rate*

Although the interest rate does not have an immediate impact on the stock market, increasing interest rate has an indirect effect to the market. If central bank decides to raise interest rate, many individual investors have to reduce their investment from stock market. To value a company, it takes the sum of all the expected future cash flow from that company discounted back to the present. All else being equal, this lowers the company's stock price. The result of our study displays the effect of changing interest rate from 1997 to 2007 to the stock composite index. The central bank had changed interest rate eleven times during that time period, and there are six times change of SHEC on the following week were opposite with interest rate change direction. This result is consistent with Mueller (2006) However, the other five times change of interest rate had positive effect to the stock price, which is inconsistent with previous reports. We also found that a big change of interest rate lead to a bigger effect to the stock price. For example, in October 1997 and June 1999, the interest rate had been decreased by 1.8 and 1.43 points respectively, and then SHEC increased 7.56% and 32.05% on the following week after announcement. On the other hand, in February 2002 and October 2004, the interest rate had changed 0.22 and 0.27 points respectively, but the SHEC only changed 2.21 and 2.75 percentages respectively. Although we can not apply GARCH class models to the interest rate effect due to the reason of only eleven changes happened which is inconsistent with stock daily data, interest rate is sensitively affect stock return either positive or negative direction in China stock market.

B. CPI

China's CPI from June 2006 to April 2007 shows that China's CPI began to increase in July 2006 and then raised quickly from October 2006 till now. At the same time, Shanghai stock index also increased from 1785 points on October 9th 2006 to 3841 points on April 30th 2007. Therefore, as CPI starting to increase, stock market also grew faster at beginning level. However, this conclusion is inconsistent with reports by Liljeblom and Stenius (1997), and Valckx(2004). These authors thought that CPI is negatively related to stock market. To see if CPI has effect to stock volatility, we apply the best in-sample model EGARCH-M (1, 1) with variance regressor of the log first difference of CPI from January 1st, 1997 to June 30th 2004. the result displays that the AIC and SC are -5.907156 and -5.883505 respectively, which is slightly smaller than the EGARCH-M(1,1) model without variance regressor. This means that the effect of CPI factor to the stock volatility is not obviously. The figure also indicate that the coefficient of CPI factor is -0.625436 with p-value 0.706, which means there exist some correlation of CPI and stock index return, but this relationship is not significant. Therefore, the inflation is stable as a macroeconomic factor, so it has limit effect to China stock market.

C. Government policy and circumstance of companies

The government policy is very important in China, since the government can use their policy lever to influence economics, industries and even the stock market. The most recently example happened in June 2005. This policy change is also an important reason why this paper separate out of sample data to two groups by the change point date of June 30th 2005. One out of sample is from July 1st 2004 up to June 30th 2005, and the other is from July 1st 2005 until April 30th 2007. Back to June 6th, 2005, China stock market already passed from five years bear market, Shanghai stock index dropped to 998 points on that day from 2245 points on June 15th, 2001. The government decided to permit most nation-owned or partly nation-owned companies to give their non-trading shares to the existing trading shareholders. Actually, this policy was recognized by increasing supply of stocks, so that it may cause the further drop of stock market. However, this change makes most companies stock fully exchange in the stock market, so it encourages company to improve their management and payoff level. The companies now can control their capital based on the market and they are totally independent by national government, province government or city government. This policy recovers investors' confidence, and attracts a lot of capital from investment fund, insurance capital, pension fund and foreign financial institutions. Therefore, more and more capital entered into China stock market, which made China stock market turn to bull market until today

D. Influencing factors summary

Based on the analysis of some macroeconomics factors, we found the sensitively of stock returns to the announcement of interest rate change during the sample period with both positive of negative directions. The CPI in China does not have obviously effect to the stock market, but there still exist some correlation between inflation and stock return.The government policy is very important in China; it may guide stock market to turn into the other way.

VII. CONCLUSIONS

After trying above five GARCH class models within different in-sample periods, we found that Shanghai stock market does have some asymmetric effect within the whole sample periods. EGARCH-M (1, 1) models may be best explained return series based on five models' AIC and SC values. The forecast ability of the models in the out of sample periods is different with in-sample periods. No one single forecast model is found to be the best for Shanghai stock market. Based on ME, MAE and RMSE statistic measures, the asymmetric models, EGARCH (1, 1) and TARCH-M (1, 1), are the best models during the out of sample periods from July 1st, 2004 to June 30th, 2005. Alternatively, for a longer and rapid growth period from July 1st 2005 to April 30th 2007, a simple GARCH (1, 1) model was recommended to the best model by more reliable error measurement MAE and RMSE.

Although most evidences in empirical finance indicate that GARCH class models are good for forecasting returns on financial assts, we ought to try some other models to see if they are better fitted to sample data. Some multivariate GARCH models allow time-varying variance-covariance have been introduced in the recent literature. Some non-linear GARCH models can explain skew distribution of the return, like Quadratic GARCH model and GJR model; they also have been investigated by some researchers.

REFERENCES

[1] Balaban E. and Bayar A. (2006) Forecasting Stock Market Volatility: Further International Evidence The European Journal of Finance Vol12, 2,171-188

[2] Brooks R. D. and Ragunathan Vanitha (2003) Returns and volatility on the Chinese stock markets Applid Financial Economics 13, 747-752

[3] Hol, E.M.J.H. (2003) Empirical Studies on Volatility in International Stock Markets. Boston: Kluwer Academic Publishers.

[4] Knight, J.(2007) Forecasting Volatility in the Financial Markets Third Edition. Oxford: Elsevier

[5] Li H. (2007) International Linkages of the Chinese Stock Exchanges: a multivariate GARCH analysis Applid Financial Economics 17, 285-297

[6] Pan H. and Zhang Z. (2006) Forecasting Financial Volatility: evidence from Chinese stock market Ph.D. thesis, University of Durham

This paper supported by the Fundamental Research Funds for the Central Universitie , No RW2010-3.