An Investigation of Conditional Heteroscedasticity Structural Change in S&P 500 Returns

Jinyu Du[1], V.A. Samaranayake[2]

[1,2]Missouri University of Science and Technology, 400 West 12th Street, Rolla, MO 65409

Abstract

It is well-known that financial returns such as those obtained from S&P 500 index data exhibits conditional heteroscedasticity. In this study, we investigate whether the conditional volatility structures of S&P500 differ between economic recession and non-recession periods. This investigation was performed on S&P500 returns from1989 – 2015 as well as S&P 500 sector returns from 2007 – 2017. In initial investigations into an appropriate volatility model for this data, the EGARCH (1,1) was found to be the optimal model, indicating the underlying asymmetric conditional volatility structure caused by positive and negative shocks of news/innovation. Regression analysis on the logarithm of squared return data was performed to determine whether model parameters changed across different time segments. News Impact Curves (NIC) were plotted to visualize the underlying differences among models for the selected time periods. In general, negative news/shocks induced higher conditional volatility change than positive news/shocks. Results indicated that volatility structures during the non-recession periods are significantly different from those of the recession periods, with the latter inducing more volatility. S&P 500 sector returns also showed similar patterns.

Keywords: Conditional Heteroscedasticity, GARCH Models, Volatility Structure, News Impact Curves, Regression Analysis.

1. Introduction

Financial returns, including those obtained from S&P 500 index data, exhibits conditional heteroscedasticity with common characteristics such as stationarity (at least locally), volatility clustering, and leverage effects. The EGARCH model of Nelson (1991), TGARCH model of Glosten (1993) and PGARCH of Ding (1994) are typical non-linear models that are good at capturing the above properties. Asymmetric GARCH models have been shown to be superior at predicting the volatility of the S&P 500 stock index (Awartani, 2005). A study on the Chinese stock market (Chang, S. 2010) has shown that long term volatility is greater during financial crisis periods. Bad news produced stronger effect than good news in the Chinese stock market during the criss. This motivated us to ask the question: how did the volatility of the U.S. stock market behave before and after the 2008 financial crisis as well as other economic recession periods? Past studies have shown evidence of structural breaks in the S&P 500 time series (Chu, 1995; Smith, D.R. 2008) in periods prior to 2008. Euro exchange rate return over the 1999-2013 period also showed structural breaks in the unconditional variance and structural GARCH models outperform non-structural GARCH (Trenca, I., 2015) during this time span. Multiple structural breaks in the conditional variance dynamics of asset returns associated with the Asian and Russian financial crises were also found by Andreou (2002).

2. The EGARCH Model

The Exponential GARCH (EGARCH) model was proposed by Nelson (1991) and accounts for the possibility of differential impact of positive and negative shocks on conditional volatility. It models the conditional variance σ_t^2 of the returns (or log returns) ε_t as follows:

$$\varepsilon_t = \sigma_t z_t, \quad \text{with } \{z_t\} \sim i.i.d.\ (0,1), \qquad (1)$$

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{q} \alpha_i g\left(z_{t-i}\right) + \sum_{j=1}^{q} \beta_j \ln(\sigma_{t-j}^2),$$

where

$$g(z_t) = \theta * (z_t) + \left[\gamma |z_t| - E|z_t|\,\right]. \qquad (2)$$

Note that in SAS AUTOREG procedure, which was used in this analysis, the coefficient γ in $g(z_t)$ is set to one. Observe also that $E|z_t| = \sqrt{2/\pi}$ if $z_t \sim$ N (0, 1).

Based on results in Reider (2009) that expresses the log of conditional volatility in the form of an AR process, with AR coefficients related to the EGARCH coefficients, we developed a model to test for parameter changes in the EGARCH model using regression analysis. Since in this study an EGARCH(1, 1) model was considered, the corresponding autoregressive model for the series $\{\sigma_t^2\}$ is given by

$$\ln\left(\sigma_t^2\right) = \eta_0 + \eta\left[\ln\left(\sigma_{t-1}^2\right) - \eta_0\right] + g\left(z_{t-1}\right)$$
$$= \mathrm{a} + \mathrm{b}\ln\left(\sigma_{t-1}^2\right) + \delta_{t-1},$$

where g is defined as in (2). Notice that it is common practice to use square of the log returns to approximate the values of the conditional volatility. Thus we used the values $S_t = (\varepsilon_t - \mu)^2$ and the $S_{t-1} = (\varepsilon_{t-1} - \mu)^2$, where μ is the overall mean of the log returns, to obtained estimates of the conditional volatilities σ_t^2 and σ_{t-1}^2 respectively.

Three recession periods were identified over the time span of the observed S&P 500 data used in this study. They are: July 2, 1990 – March 28, 1991; March 1, 2001, November 30, 2001; December 3 2007 – June 30, 2009 (see Wikipedia article: list of recessions in the United States). This results in seven periods of non-recession and recession time segments. Dummy variables were created for six of the recession and non-recession periods, excluding the period prior to the first recession. The dummy variables and their interaction with the lag conditional volatility term produced the following regression model:

$$\ln(S_t) = \mathrm{a} + b * \ln(S_{t-1}) + D_1 + D_2 + \ldots + D_6 + c_1 * D_1 * \ln(S_{t-1}) + \ldots$$
$$c_2 * D_2 * \ln(S_{t-1}) + \ldots + c_6 * D_6 * \ln(S_{t-1})\,. \qquad (3)$$

The above model was then fitted to the mean adjusted log return. The mean of the original log returns, while small, was subtracted in order for the data to better fit the assumed underlying model.

3. Standard and Poor's 500 Stock Index Data

The Standard and Poor's 500 (S&P 500) index includes 500 leading companies and captures approximately 80% coverage of available market capitalization. It is widely accepted as the best single gauge of large-cap U.S. equities. Companies covered by the index can are generally categorized into eleven industrial sectors: (1) Information Technology, (2) Health Care, (3) Financials, (4) Consumer Discretionary, (5) Industrials, (6) Consumer Staples, (7) Energy, (8) Utilities, (9) Real Estate, (10) Materials, and (11) Telecommunication services. Detailed sector industry information is provided in Table 1.

3.1 Source of the Data and Data Description

The S&P 500 index data was downloaded from Yahoo Finance (see website). The S&P 500 sector data was obtained from S&P Dow Jones Indices website. For the S&P 500 index data, we selected data from January 3, 1989 to December 31, 2015. The S&P 500 sector data, from June 29, 2007 to July 14, 2017, were selected to cover the time span covered the recession period of interest. Definitions of sectors are given in Table 1.

3.2 Pre-Processing of the Data

Daily total returns P_t was selected for the analysis. Based on total returns, continuously compounded returns (also known as log returns), $r_t = \ln\left(P_t / P_{t-1}\right)$, were computed for analysis. The reason for using log returns is that r_t enjoys some advantages over simple net returns $r_t = \ln\left(P_t / P_{t-1}\right) - 1$ and its statistical properties are more tractable (Tsay, R.S. 2012). Note that r_t is our observed value of the returns ε_t given in model (1). The graph of the 27-year log return data is given in Figure 1. The plot as well as the test results indicates the presence of conditional heteroscedasticity in the returns data.

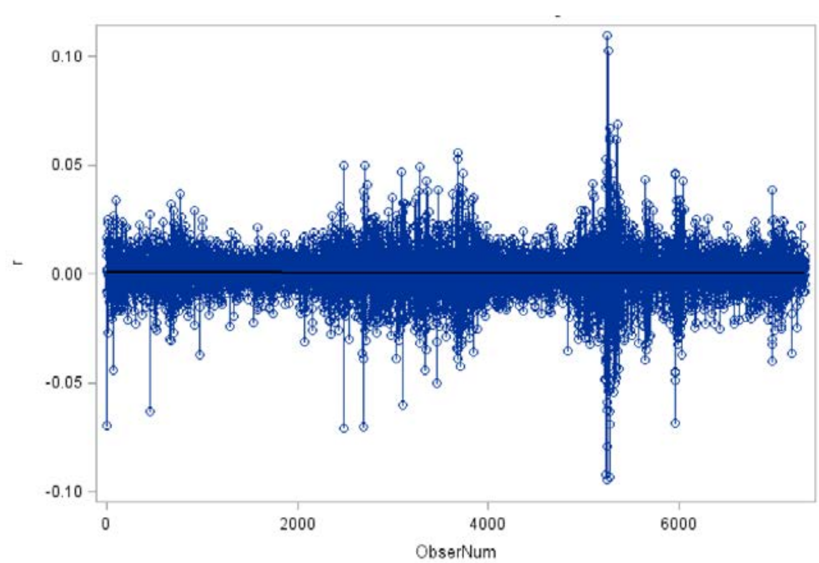

Figure 1. S&P 500 log Return (1989 – 2015) Results

Table 1. Industries Belong to different S&P 500 Sectors

Sector	**Industries**
Information Technology	Internet Software & Services, IT Services, Software, Communications Equipment, Computers & Peripherals, Electronic Equipment & Components, Office Electronics, Semiconductor Equipment and Products (discontinued effective 04/30/2003), Semiconductors & Semiconductor Equipment
Health Care	Healthcare Providers & Services, Healthcare Equipment & Supplies, Healthcare Technology, Biotechnology, Pharmaceuticals, Life Sciences Tools & Services
Financials	Commercial Banks, Thrift & Mortgage Finance, Diversified Financial Services, Consumer Finance, Capital Markets, Insurance, Real Estate (discontinued effective 04/30/2006), Real Estate Investment Trusts, Real Estate Management & Development
Consumer Discretionary	Auto Components, Automobiles, Household Durables, Leisure Equipment & Products, Textiles Apparel & Luxury Goods, Hotels, Restaurants & Leisure, Diversified Consumer Services, Media, Distributors, Internet and catalog Retail, Multiline Retail, Specialty Retail
Industrials	Aerospace & Defense, Building Products, Construction & Engineering, Electrical Equipment, Industrial Conglomerates, Machinery, Trading Companies & Distributors, Commercial Services & Supplies, Professional Services, Air Freight & Logistics, Airlines, Marine, Road & Rail, Transportation Infrastructure
Consumer Staples	Food staples and Retailing, Beverages, Food Products, Tobacco, Household Products, Personal Products
Energy	Energy Equipment & Services, Oil, Gas, & Consumable Fuels
Utilities	Electric Utilities, Gas Utilities, Multi-Utilities, Water Utilities, Independent Power Producers & Energy Traders
Materials	Chemicals, Construction Materials, Containers & Packaging, Metals & Mining, Paper & Forest Products
Telecommunication services	Diversified Telecommunication Services, Wireless Telecommunication Services

4. Modeling Procedure and Results

4.1 Modeling Procedure

A preliminary analysis was carried out to determine which of GARCH, EGARCH, PGARCH, TGARCH, IGARCH, as well as additional autoregressive components such as AR-1, AR-2 would best fit the data. The SAS AUTOREG procedure was employed to carry out this model selection, with the Akaike Information Criteria (AIC) employed to compare the models. Once a model was identified, the AR model given in equation (3) was fitted to the S&P 500 mean adjusted log return data to identify changes in the model parameters across the seven recession/non-recession periods. Then separate models were fitted to each segment that was identified as different and the estimated parameters values were employed to draw the News Impact Curves (NIC).

4.2 Results

Initial model selection identified EGARCH as the overall best model to describe the S&P 500 log return data. The results of tests for parameter changes in the EGARCH model using regression analysis are listed in Figure 2. It shows that the parameters for recession period are significant, either on their own or in an interaction term; parameters for non-recession period are insignificant and the interaction between recession period 1 and the lag conditional variance term is significant. All other interactions terms were not significant. This indicates that while the recession periods are distinct between themselves and the non-recession periods, the non-recession periods were not statistically different from one another as far as volatility propagations is concerned. One caveat is that Model (3) only looks at changes in the intercept parameter of the EGARCH model and assumes that the rest of the parameters in the model remain the same across different periods. As we see from the individual fits of the EGARCH model to different data segments, even the non-recession periods appears different from one another.

The NIC curves for EGARCH model fitted to the S&P 500 27-year log return data is shown in Figure 3. It shows that negative news causes a much impact on stock price than the positive news. Compared to Recession Periods 1 and 2, Recession Period 3 shows more sensitivity to negative news. This sensitivity to negative news seems even more pronounced after the end of Recession 3. The NIC curves for S&P 500 10-year sector log return data, with respect to Recession Period 3 is shown in Figure 4. It is seen that during the 3rd recession period, consumer staples sector, health care sector and utilities sector reacted stronger to negative innovation/news compared to the rest of the other sectors. The NIC curve for EGARCH model for non-recession periods of S&P 500 10-year sector log return data is shown in Figure 5. It is seen that during the period after Recession 3, consumer discretionary sector, financials sector and industrials sector reacted stronger than the other sectors to negative news. Overall, all sectors react to negative innovation/news much more than the positive ones.

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1631.61396	326.32279	52.41	<.0001
Error	7311	45523	6.22659		
Corrected Total	7316	47154			

	Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
	Intercept	-10.28405	0.13398	36688	5892.14	<.0001
*	LSrL	0.07452	0.01165	254.64451	40.90	<.0001
	D3	0.89572	0.18384	147.81723	23.74	<.0001
	D4	0.14286	0.07323	23.70036	3.81	0.0511
	D5	1.66317	0.13172	992.73645	159.43	<.0001
	D1LSrL	-0.06686	0.01751	90.82064	14.59	0.0001
* Forced into the model by the INCLUDE= option						

Figure 2. Regression Testing for EGARCH(1,1) Model Parameter Changes

Notation: LSrL= ln(S_{t-1}), D1- Recession 1, D2- period between Recession 1 and 2, D3-Recession 2, D4- period between Recession 2 and 3, D5 – Recession 3, D6 – period after Recession 3. Dummy variables not shown are insignificant. When D1 was forced in to make a hierarchical model, D1LSrL became insignificant while D1 became significant.

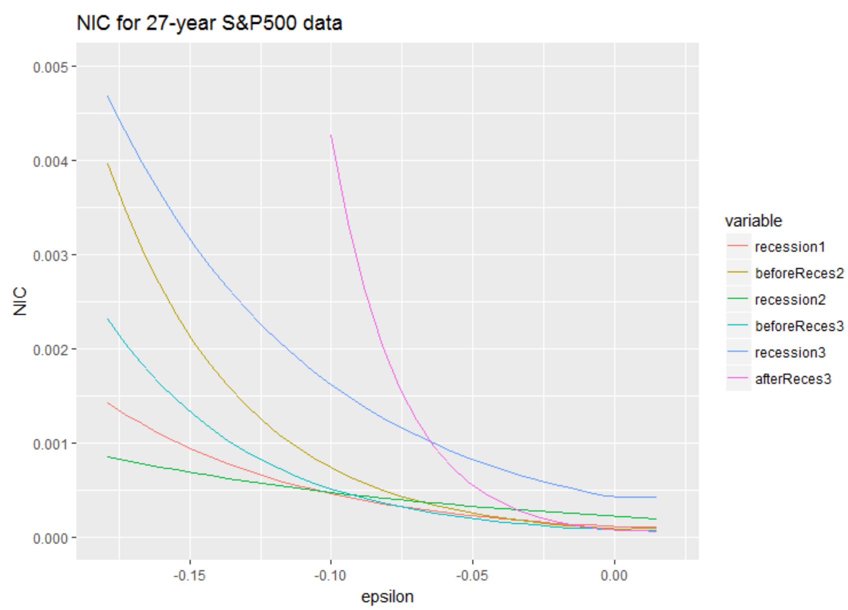

Figure 3. NIC curve for EGARCH Model (S&P 500 Log Return 27-year Data)

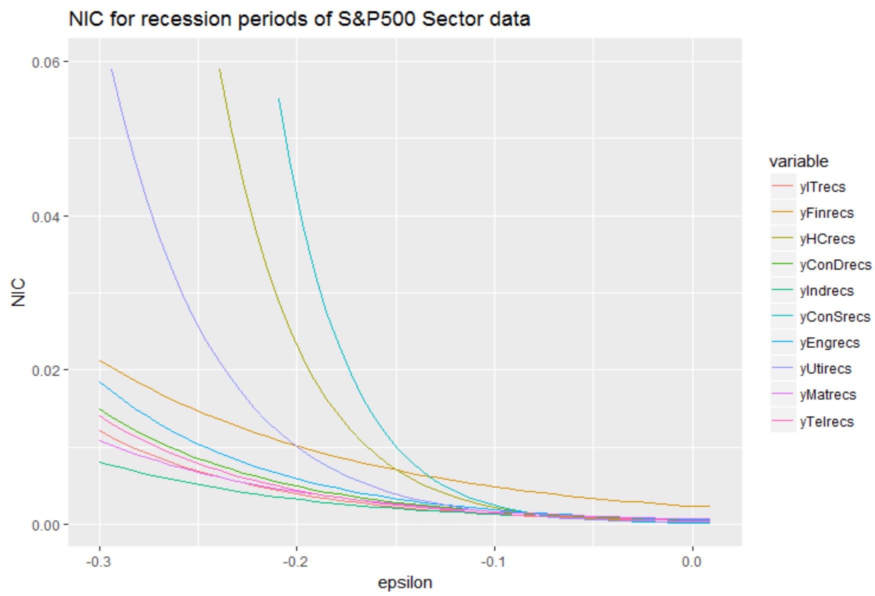

Figure 4. NIC curve for Recession Periods EGARCH Model (S&P 500 Sector 10-year Data)

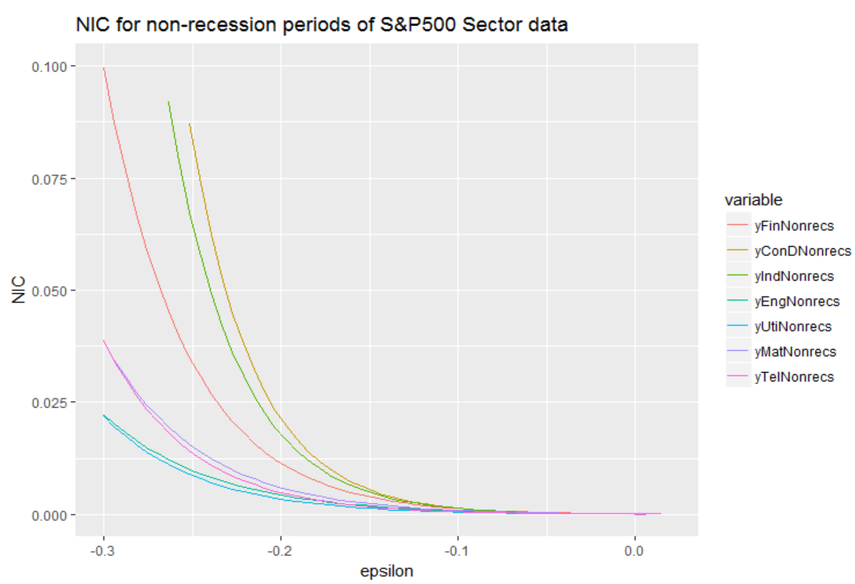

Figure 5. NIC curve for Non-recession Periods EGARCH Model (S&P 500 Sector 10-year Data)

5. Conclusions

The modeling and analysis shows that non-symmetric EGARCH(1,1) model is the best choice for modeling return and conditional volatility of S&P 500 return data as well as individual sector data. It also shows that the volatility structure during economic recession periods and non-recession periods are statistically different.

News Impact Curves (NIC) show that for all time segments, negative news impact the conditional volatility far more than positive news, but the NIC changes across recessions and non-recession periods as well as across sectors. Detailed sector data analysis shows that during the 2007 – 2009 recession period, negative news impacted Consumer Staples, Health Care and Utility sector more than other sectors, whereas during the non-recession period, negative news impacted Consumer Staples, Health Care and Information Technology sector more than other sectors.

6. Acknowledgements

The authors would like to thank the department of Mathematics & Statistics at Missouri S&T for the research facilities and funding that enabled the successful conduct of this research.

References

Andreou, E. and Ghysels, E. (2002) Detecting multiple breaks in financial market volatility dynamics. Journal Of Applied Econometrics, 17, 579–600.

Awartani, B.M.A. and Corradi, V. (2005) Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries. International Journal of Forecasting, 21, 167– 183.

Chang, S. (2010) Application of EGARCH model to estimate financial volatility of daily returns: The empirical case of China. University of Gothenburg: Master Degree Project No.2010:142.

Chu, C. & Shang J., (1995) Detecting parameter shift in GARCH models. Econometric Reviews, 14, 241-266.

Ding, Z., Granger, W.J. and Engle, R.F. (1993) A long memory property of stock market returns and a new model. Journal of Empirical Finance, 1:83–106.

Engle, R.F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. Inflation. Econometrica, 50, 987-1008.

Engle, R.F., & Ng, V. (1993) Measuring and testing the impact of news on volatility. The journal of finance, 48(5), 1749-1778. doi:10.2307/2329066.

Franses, P. H. and Van Dijk. D., (2000) Non-linear time series models in empirical finance. United Kingdom, UK: University Press, Cambridge.

Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance, 48: 1779–1801. doi:10.1111/j.1540-6261.1993.tb05128.x

List of recessions in the United States (n.d.). In *Wikipedia*. Retrieved December 1, 2016, from https://en.wikipedia.org/wiki/List_of_recessions_in_the_United_States.

Nelson, D. (1991) Conditional heteroscedasticity in asset returns: A new approach. Econometrica, 59(2), 347-370. doi:10.2307/2938260.

Pagan, A.R. and G.W. Schwert, (1990) Alternative models for conditional stock volatility. Journal of Econometrics, 45, 267-90.

Reider, R. (2009) Volatility forecasting I: GARCH models. Retrieved from http://cims.nyu.edu/~almgren/timeseries/Vol_Forecast1.pdf.

Samaranayake VA and Liu, J. 2015. An Investigation of the Day-of-the-Week Effect on the Volatility and Returns of Individual S&P 500 Sectors. In JSM Proceedings, Business and Economic Statistics Section. Alexandria, VA: American Statistical Association. 2991–3001.

Smith, D. R. (2008) Testing for structural breaks in GARCH models. Applied Financial Economics, 18(10), 845-862.

S&P Dow Jones Indices. (2017). S&P 500. Retrieved from http://us.spindices.com/indices/equity/sp-500

Trenca, I., Zapodeanu, D. and Cociuba, M. I. Testing the presence of structural break in the euro exchange rate variance (2015) Procedia Economics and Finance, 32 (2015) 1163 – 1169.

Tsay, R.S. (2012) An introduction to analysis of financial data with R. United states. Hoboken: John Wiley and Sons Ltd.

Yahoo Finance. (2016). S&P 500 SNP - SNP Real Time Price. Retrieved from https://finance.yahoo.com/quote/%5EGSPC/history?period1=852012000&period2=1483164000&interval=1d&filter=history&frequency=1d