VALUE-AT-RISK MEASURES DURING CRISIS PERIODS

by

Karl Kuppers
BSc. Chemistry, Simon Fraser University, 2012

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN FINANCE

In the Master of Science in Finance Program
of the
Faculty
of
Business Administration

SIMON FRASER UNIVERSITY

Fall 2018

Approval

Name: **Karl Kuppers**

Degree: **Master of Science in Finance**

Title of Project: **Value-At-Risk Measures During Crisis Periods**

Supervisory Committee:

____

Andrey Pavlov, Ph.D.
Senior Supervisor
Professor of Finance

____

Phil Goddard, Ph.D.
Second Reader
Instructor

Date Approved: ______________________________

Abstract

Value-at-Risk (VaR) is a commonly used measure of market risk in the financial industry. The measure seeks to identify the loss that an investment may realize within a certain confidence level. It is used to evaluate the market risks of assets and to calculate capital reserve requirements. However, despite its wide use in financial risk management, it has several well-known limitations. This project analyses a class of VaR models that were published in the academic literature and evaluates their performance when market conditions are stressed. The findings of this research reveal that these models perform poorly when market conditions are changing. Managers who rely on these VaR models may underestimate their risks and fail to set aside appropriate capital reserves to handle adverse market moves.

Keywords: value-at-risk; market risk; crisis; risk management; GARCH

Acknowledgements

First, I would like to thank Andrey Pavlov for his supervision of this project and his suggestions and encouragement throughout the thesis process. I would also like to thank Phil Goddard for his suggestions and comments on this project. Finally, I would like to thank Robert Engle and Simone Manganelli for publishing an excellent article on Value-at-Risk models and for making the source code and data for their publication available on the web to aid in the replication and understanding of their research.

Table of Contents

List of Figures

List of Tables

Glossary

VaR	Value-at-Risk
DQ	Dynamic Quantile
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
CL	Confidence Level
OOS	Out-of-sample
CAViaR	Conditional Autoregressive Value-at-Risk

1. Introduction

Value-at-Risk (VaR) is a commonly used risk measure that financial institutions and supervisors use to measure market risk. It attempts to calculate the amount that a financial institution can lose over a given time period within a specified confidence level. The output of a VaR estimate enables a manager to say something like "99% of the time we will lose less than $10,000 in one day".

Due to its usefulness in providing a bank's management with an understanding of its market risk, VaR has been used by financial institutions since the 1980s. In 1995, the Basel Committee of Banking Supervision published an amendment to Basel I that encouraged financial institutions to use VaR to calculate their capital reserve requirements for market risk (BIS, 1995). Financial institutions currently use it to help determine the riskiness of various investments, and this helps their managers decide if they should pursue those investments.

The following is a framework to calculate VaR (Engle & Manganelli, 2004):

1. The price of the portfolio and its components is revalued daily (marked-to-market) and these prices are stored in a database
2. From the database of portfolio prices, a distribution of portfolio returns is generated
3. From the distribution of returns, the VaR of the portfolio is calculated at a specified quantile

Although VaR models commonly follow this framework, there are differences in how each step is carried out. For the first point, there are questions regarding how long the database of prices should be. A short time series may not sufficiently capture the effects of the market, but a long time series may have characteristics that are no longer relevant. There may also be concerns about the quality of the time series data, especially if the portfolio is composed of illiquid assets that may be difficult to value daily. For the second point, there are questions about what kind of a distribution to use, because different distributions have been found to be more useful under different market conditions. Additionally, correlations between assets have been found to vary over time, particularly during crisis periods, and this can impact the validity of the distribution. For the third point, there are questions over how high the confidence level should be set in order to adequately capture the risks, especially in light of recent economic crises, where many banks were unprepared for the severity of the effects.

Despite the attractiveness and popularity of VaR, it suffers from several shortcomings. While it specifies the maximum that a manager expects to lose at a given confidence level, it says nothing about what may happen if extreme events do occur. For example, while the 99% VaR is an estimation of the maximum loss that will be experienced 99% of the time, it says nothing about what the loss will be should an event in the other 1% occur, other than that it will be greater than the 99% VaR. Another criticism of VaR is that it is not sub-additive. Research done by Artzner et al. (1999) demonstrates that the risk of combining two assets individually may be greater than the risk of a portfolio that includes both assets. Their findings also suggest that, under certain circumstances, VaR opposes the benefits of diversification and favours the concentration of assets, because diversification can increase the value-at-risk.

This project will investigate several VaR models that were published in the academic literature and compare and evaluate the performance of those models by applying them to market conditions at the time those papers were published as well as the market conditions of the economic crisis in 2008-2009.

2. Background

In 2004, Robert Engle and Simone Manganelli published a paper titled "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles" where they introduced a way to calculate VaR by modelling the quantile of interest directly rather than modelling the whole distribution and extracting the VaR from that distribution (Engle & Manganelli, 2004). They believed that since the volatility of stock returns was observed to exhibit autocorrelation and cluster over time, the VaR should have similar behaviour, and thus it should be possible to directly model the VaR quantile through an autocorrelation model. Generally, these models would take the following form:

$$\text{VaR}_t = \beta_0 + \sum_{n=1}^{q} \beta_n \text{VaR}_{t-n} + \sum_{m=1}^{r} \gamma_m f(\boldsymbol{x}_{t-m})$$

where VaR_t is the risk measure at time *t*, VaR_{t-n} is the risk measure at time *t* - *n*, $f(\boldsymbol{x}_{t-m})$ is the value of a function applied to variables $\boldsymbol{x}$ that were observable at time *t* – *m*, and the βs and γs are the model parameters to be estimated.

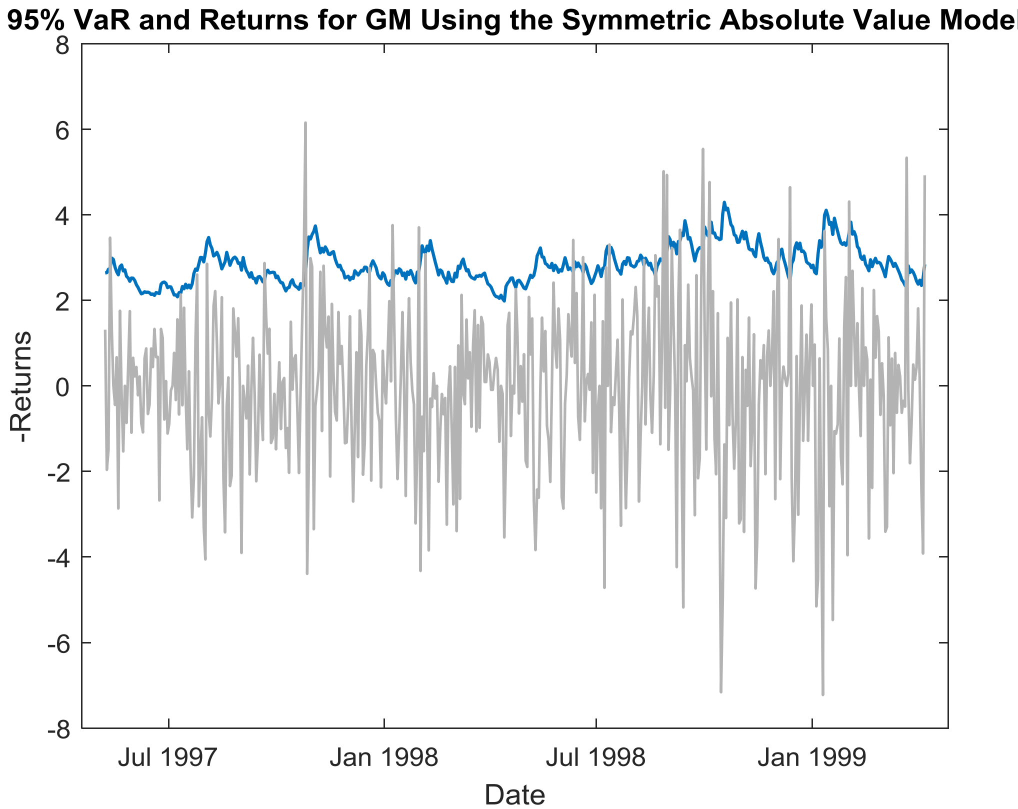
The authors proposed the following four VaR models in their paper, which will be investigated further in this research project:

Symmetric Absolute Value:

$$\text{VaR}_t = \beta_1 + \beta_2 \text{VaR}_{t-1} + \beta_3 |y_{t-1}|$$

In this model, y_{t-1} is the value of the return at time *t* – 1. This model specifies the VaR at time *t* as being a combination of a constant, the VaR at time *t* – 1, and the absolute value of the actual return at time $t - 1$. A visual representation of this model is presented in Figure 1.

Figure 1 *VaR and Returns Using the Symmetric Absolute Value Model*

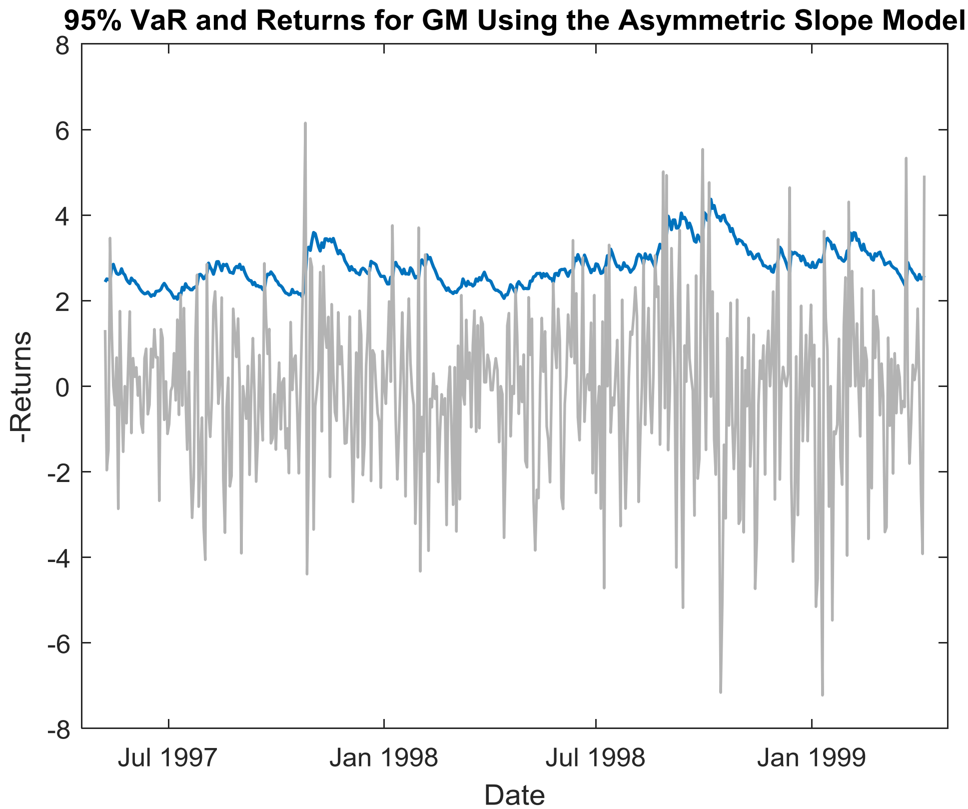


Asymmetric Slope:

$$\mathrm{VaR}_t = \beta_1 + \beta_2 \mathrm{VaR}_{t-1} + \beta_3 \max(y_{t-1}, 0) - \beta_4 \min(y_{t-1}, 0)$$

This model is similar to the Symmetric Absolute Value model, except it also allows for positive and negative returns at time $t - 1$ to be given different weightings. A visual representation of this model is presented in Figure 2.

Figure 2 *VaR and Returns Using the Asymmetric Slope Model*

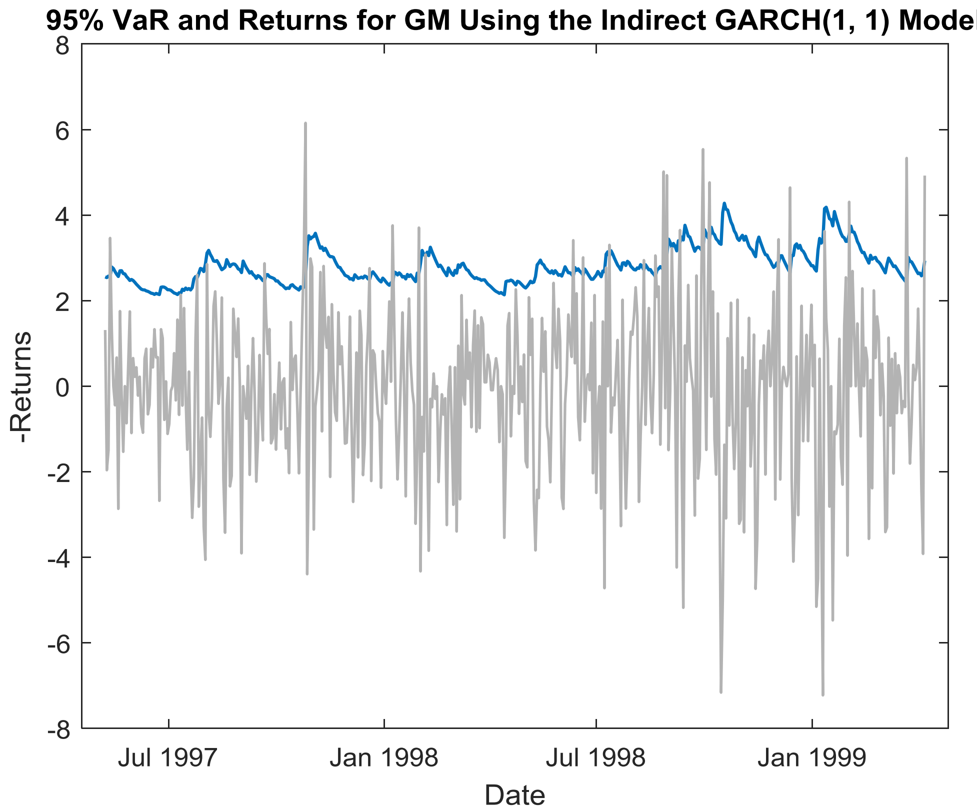


Indirect GARCH(1, 1):

$$\text{VaR}_t = \sqrt{\beta_1 + \beta_2 \text{VaR}_{t-1}^2 + \beta_3 y_{t-1}^2}$$

The authors referred to this model as "Indirect GARCH" because it is unclear if the true data generating process is an actual GARCH process with independent and identically distributed (iid) error terms. The authors stated that the model is not as restrictive as GARCH in this way, and their model will still work if the error densities and volatilities are not constant. A visual representation of this model is presented in Figure 3.

Figure 3 *VaR and Returns Using the Indirect GARCH(1, 1) Model*

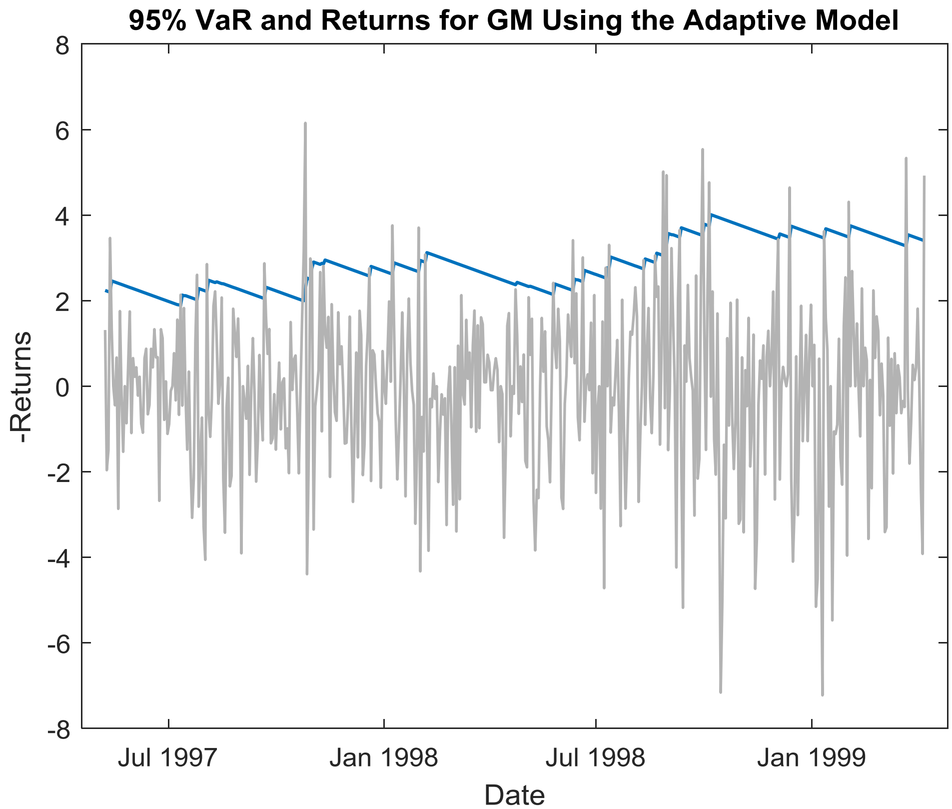


Adaptive:

$$\mathrm{VaR}_t = \mathrm{VaR}_{t-1} + \beta_1[(1 + \exp[G(y_{t-1} - \mathrm{VaR}_{t-1})])^{-1} - \theta]$$

In this model, G is defined as a positive finite constant and θ is the VaR quantile. For example, if we are modelling the 95% VaR, θ would be specified as 0.05. This model computes the VaR at time t by looking at the VaR at time $t - 1$ and seeing if the actual return exceeded it. If the actual return did exceed it, the VaR estimate for time t is increased by a fixed amount, like a step function. If the VaR estimate for time $t - 1$ was not exceeded by the actual $t - 1$ return, the VaR estimate for time t is decreased slightly by a fixed amount. A visual representation of this model is presented in Figure 4.

Figure 4 *VaR and Returns Using the Adaptive Model*



To estimate the model parameters $\boldsymbol{\beta}$, the authors used the following regression quantile function (Koenker & Bassett Jr, 1978):

$$\min_{\boldsymbol{\beta}} \frac{1}{T}\sum_{t=1}^{T}[\theta - I(y_t < \mathrm{VaR}_t(\boldsymbol{\beta}))][y_t - \mathrm{VaR}_t(\boldsymbol{\beta})]$$

where T is the number of samples and I is a function such that $I(y_t < \mathrm{VaR}_t(\boldsymbol{\beta}))$ is a vector of binary variables that are 0 and 1 depending on if the return exceeded the VaR level or not.

Some of the more recent work that was published on these models is as follows. In 2006, Kuester et al. proposed a CAViaR model that allowed for a conditional mean that varies in time (Kuester, Mittnik, & Paolella, 2006). In 2008, Taylor presented a way to use these models to calculate Expected Shortfall, which is the expected return when the VaR level has been exceeded (Taylor, 2008). In 2009, Huang et al. proposed a modification to the Asymmetric Slope CAViaR model that allows for it to be estimated with three parameters instead of four (Huang, Yu, Fabozzi, & Fukushima, 2009). In 2013, Jeon and Taylor proposed CAViaR models that included a term for implied volatility (Jeon & Taylor, 2013).

This research project will apply the original models to more recent market data and critically evaluate their suitability as measures of market risk.

3. Methodology

The original research done by Engle and Manganelli (2004) used daily market return data for General Motors (GM), IBM, and the S&P 500 Index from April 7, 1986 to April 7, 1999. The authors obtained 3,392 daily returns, with 2,892 used for in-sample model parameter estimation and the remaining 500 for out-of-sample evaluation. This paper replicates the results from the original data and considers the following new data. The same securities were chosen, with dates ranging from January 23, 1996 to January 22, 2009. It was observed that the original data contained returns on holidays, so holidays were included in the new data as well. The new data again provides 2,892 data points for in-sample estimation and 500 points for out-of-sample evaluation. All market data for this research was obtained from the Bloomberg Terminal (Bloomberg, 2018). As GM declared bankruptcy in 2009, the ticker chosen for GM was Motors Liquidation Company (MTLQQ). The tickers for IBM and the S&P 500 Index were IBM and SPX, respectively.

The VaR models were estimated using the same methodology as the original paper by Engle and Manganelli (2004). The authors provided their original MATLAB code and data for their paper through their website, which was utilized in this project to replicate their results (Manganelli, 2018). Unfortunately, as it was not possible to obtain the original version of MATLAB 6.1 that the authors used in their paper, this project utilizes MATLAB R2018b. The newer version of MATLAB necessitated a few changes in the code because certain MATLAB optimization algorithms were removed recently, and these algorithms had been used in the original code.

To evaluate the models for their suitability, two tests were examined in this paper. One test is the Bernoulli coverage test which tests to see if the number of VaR exceedances is significantly different from what we would expect (Daníelsson, 2011). This test is defined as the following:

$$2\log\frac{(1-\hat{p})^{\nu_0}(\hat{p})^{\nu_1}}{(1-p)^{\nu_0}(p)^{\nu_1}} \sim \chi^2_{(1)}$$

where $\hat{p}$ is the estimated probability of a VaR exceedance, p is the actual probability, ν_0 is the number of times that VaR is not exceeded, and ν_1 is the number of times that VaR is exceeded. The test statistic is then compared with the chi-squared distribution with one degree of freedom to see if the frequency of VaR exceedances is different from what is expected.

While the Bernoulli coverage test indicates if the model predicts an appropriate number of VaR exceedances, it says nothing about if those exceedances are independent or if they are clustered. A suitable VaR model should adapt fast enough such that it does not have clusters of VaR exceedances (Engle & Manganelli, 2004), so another test that could be utilized is one that tests if the VaR exceedances exhibit serial correlation. Engle and Manganelli (2004) introduced the following test for this:

$$\frac{\mathbf{Hit}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Hit}}{\theta(1-\theta)} \sim \chi^2_q$$

where **Hit** is a vector of binary variables indicating if the VaR was exceeded or not and **X** is a *T* by *q* matrix where *T* is the number of data points in the time series. The columns of **X** consist of a regressor constant (ones), the forecasted VaR, and lagged values of **Hit** indicating if there are previous VaR exceedances, with one column for each lagged value.

The authors show how this test statistic follows a chi-squared distribution with *q* degrees of freedom and tests the VaR model to see if it is unbiased, has an appropriate number of exceedances, and ensures that the exceedances are independent both of themselves and of the VaR estimate.

4. Results and Discussion

4.1 Model Estimation Parameters

The results of the model estimations from the 1986 – 1999 data at the 95% VaR and 99% VaR are presented in Tables 1 and 2, respectively. The tables show that β_2, the parameter assigned to the lagged VaR level, is statistically significant in all the models that use it, indicating the presence of autocorrelation in the tails of the distribution. Additionally, the in-sample Exceedances appear to be close to the levels that we expect from the VaR confidence level, indicating that the optimizer is performing well.

Table 1 Model estimation parameters and statistics for the 1986-1999 data, computed at the 95% VaR level.

	Symmetric Absolute Value			Asymmetric Slope			Indirect GARCH			Adaptive		
	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500
Beta1	0.1624	0.0521	0.0068	0.0728	0.0949	0.0381	0.3305	0.5383	0.0256	0.2785	0.3991	0.3691
p values	0.0709	0.2451	0.1741	0.0020	0.0262	0.0023	0.0005	0.0001	0.0034	0	0	0
Beta2	0.8837	0.9150	0.9619	0.9322	0.8920	0.9027	0.9047	0.8261	0.9288			
p values	0	0	0	0	0	0	0	0	0			
Beta3	0.1152	0.1319	0.0689	0.0429	0.0575	0.0368	0.1213	0.1589	0.1399			
p values	0	0	0.0472	0.0222	0.0188	0.0473	0.1397	0.0633	0.4118			
Beta4				0.1246	0.2127	0.2867						
p values				0	0	0						
Exceedances (%)	5.0138	4.9793	5.0138	5.0138	4.9793	5.0138	4.9793	5.0138	5.0484	4.8064	4.8755	4.7372

Table 2 Model estimation parameters and statistics for the 1986-1999 data, computed at the 99% VaR level.

	Symmetric Absolute Value			Asymmetric Slope			Indirect GARCH			Adaptive		
	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500
Beta1	0.2470	0.2551	0.0059	0.3872	0.0631	0.1476	1.4773	1.3171	0.2331	0.2935	0.1606	0.5563
p values	0.1387	0.2535	0.4396	0.0700	0.1821	0.0006	0.0544	0.2490	0.0307	0.0040	0.0162	0
Beta2	0.8397	0.8802	0.9580	0.7965	0.9366	0.8729	0.7829	0.8747	0.8349			
p values	0	0	0	0	0	0	0	0	0			
Beta3	0.3461	0.2161	0.1460	0.2777	0.0671	-0.0137	0.9223	0.3400	1.0586			
p values	0.0074	0.2913	0	0.0287	0.1823	0.4525	0.2288	0.0004	0.1814			
Beta4				0.4592	0.2612	0.4969						
p values				0.0083	0.0008	0.0001						
Exceedances (%)	1.0028	0.9682	1.0028	0.9682	1.0028	1.0028	0.9682	1.0028	1.0373	1.0373	1.1757	0.9336

Tables 3 and 4 contain the results of the model estimations for the data from 1996 – 2009. Once again, the tables show that β_2 is significant, indicating strong autocorrelation in the tails of the return distribution. The optimizer appears to perform well with the new data as well, as indicated by the percentages of in-sample exceedances.

Table 3 *Model estimation parameters and statistics for the 1996-2009 data, computed at the 95% VaR level*

	Symmetric Absolute Value			Asymmetric Slope			Indirect GARCH			Adaptive		
	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500
Beta1	0.0829	0.0078	0.0148	0.0680	0.0304	0.0310	0.2039	0.0038	0.0173	0.5563	0.4118	0.3166
p values	0.0032	0.1474	0.1019	0.0029	0.0428	0.0001	0.0069	0.2094	0.0874	0	0	0
Beta2	0.9127	0.9687	0.9592	0.9265	0.9335	0.9554	0.9163	0.9864	0.9579			
p values	0	0	0	0	0	0	0	0	0			
Beta3	0.1212	0.0584	0.0670	0.0662	0.0540	-0.0138	0.1490	0.0282	0.0913			
p values	0	0.0034	0	0.0015	0.0987	0.2474	0.0278	0.3962	0.2535			
Beta4				0.1403	0.1803	0.1276						
p values				0	0.0061	0.0003						
Exceedances (%)	4.9793	5.0138	4.9793	5.0138	4.9793	5.0138	5.0138	5.0138	5.0484	4.9793	4.5989	4.6335

Table 4 *Model estimation parameters and statistics for the 1996-2009 data, computed at the 99% VaR level*

	Symmetric Absolute Value			Asymmetric Slope			Indirect GARCH			Adaptive		
	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500
Beta1	0.2935	0.0724	0.0370	0.1744	0.0473	0.1195	0.2300	0.1442	0.0633	1.0478	1.1004	0.7955
p values	0.1598	0.0713	0.0148	0.1637	0.0080	0.0005	0.1277	0.1279	0.0793	0	0	0
Beta2	0.9084	0.9115	0.9458	0.9395	0.9182	0.9130	0.9685	0.9582	0.9463			
p values	0	0	0	0	0	0	0	0	0			
Beta3	0.1121	0.2558	0.1279	0.0406	0.0554	-0.0768	0.1441	0.2234	0.2631			
p values	0.0475	0.0151	0.0011	0.1957	0.0959	0.0810	0.2887	0.4707	0.2733			
Beta4				0.1356	0.4145	0.3374						
p values				0.0064	0	0.0501						
Exceedances (%)	1.0373	1.0373	1.0373	1.0028	1.0028	1.0028	1.0373	1.0028	1.0373	1.1065	0.8990	0.8645

The estimation results also show that for the Asymmetric Slope model, β_4, the parameter assigned to negative returns, is statistically significant while β_3, the parameter assigned to positive returns, is not always significant. This indicates that negative returns may have a greater contribution to VaR than positive returns.

4.2 Model Suitability Out-of-Sample Tests

Table 5 contains the data for the test statistics for the 95% VaR models when applied to the out-of-sample data from 1986 – 1999. The *p* values are used to evaluate model suitability, with *p* values of less than 0.05 indicating the model is unsuitable.

Table 5 Results of the model suitability test statistics for the 95% VaR models when used on the data from 1986 – 1999

		1986 - 1999 Data				
		Hits OOS (%)	Coverage Test (*p* values)	Accept at 5% CL?	DQ Test (*p* values)	Accept at 5% CL?
Symmetric Absolute Value	GM	4.6	0.6776	✓	0.9993	✓
	IBM	5.6	0.5455	✓	0.4447	✓
	S&P 500	5.4	0.6852	✓	0.0003	✗
Asymmetric Slope	GM	5.4	0.6852	✓	0.8892	✓
	IBM	7.4	0.0211	✗	0.0069	✗
	S&P 500	6.4	0.1678	✓	0.0007	✗
Indirect GARCH	GM	4.8	0.8364	✓	0.9014	✓
	IBM	7.2	0.0337	✗	0.1409	✓
	S&P 500	5.8	0.4229	✓	0.0001	✗
Adaptive	GM	6.0	0.3192	✓	0.3691	✓
	IBM	5.0	1.0000	✓	0.5058	✓
	S&P 500	4.6	0.6776	✓	0.0240	✗

The results show that nearly all the models produced a statistically appropriate number of exceedances in the out-of-sample data, as indicated by the high *p* values. However, when running the DQ test and testing for serial correlation among the most recent four exceedances, a few models that passed the coverage test failed the DQ test. Specifically, the S&P 500 exhibits poor performance in all models when it is subject to the DQ test.

Figure 5 shows the 95% VaR as calculated from the Indirect GARCH model and actual returns for the S&P 500 Index, with VaR exceedances indicated by red X's. Note how the number of exceedances is appropriate, but they exhibit clustering. This may explain why the model would pass the Coverage Test and fail the DQ test. For comparison, Figure 6 shows the 95% VaR and returns for GM, also calculated by the Indirect GARCH model. Note how the number of exceedances is also appropriate, but they exhibit less clustering and are more spread out. This may explain why the model would pass both the Coverage and DQ tests.

Table 6 contains the data for the model suitability test statistics for the 99% VaR models when applied to the out-of-sample data from 1986 – 1999.

Figure 5 95% VaR and Returns for the S&P 500 Index when the Indirect GARCH Model is applied. Red X's indicate where the VaR level was exceeded.

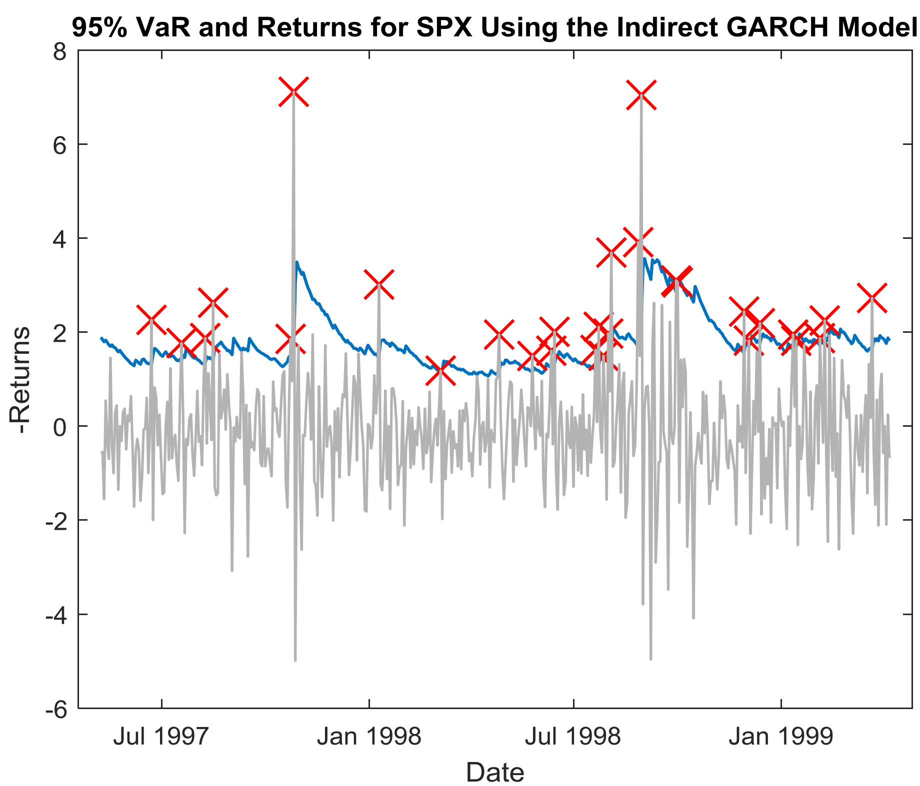


Figure 6 *95% VaR and Returns for GM when the Indirect GARCH Model is applied.*

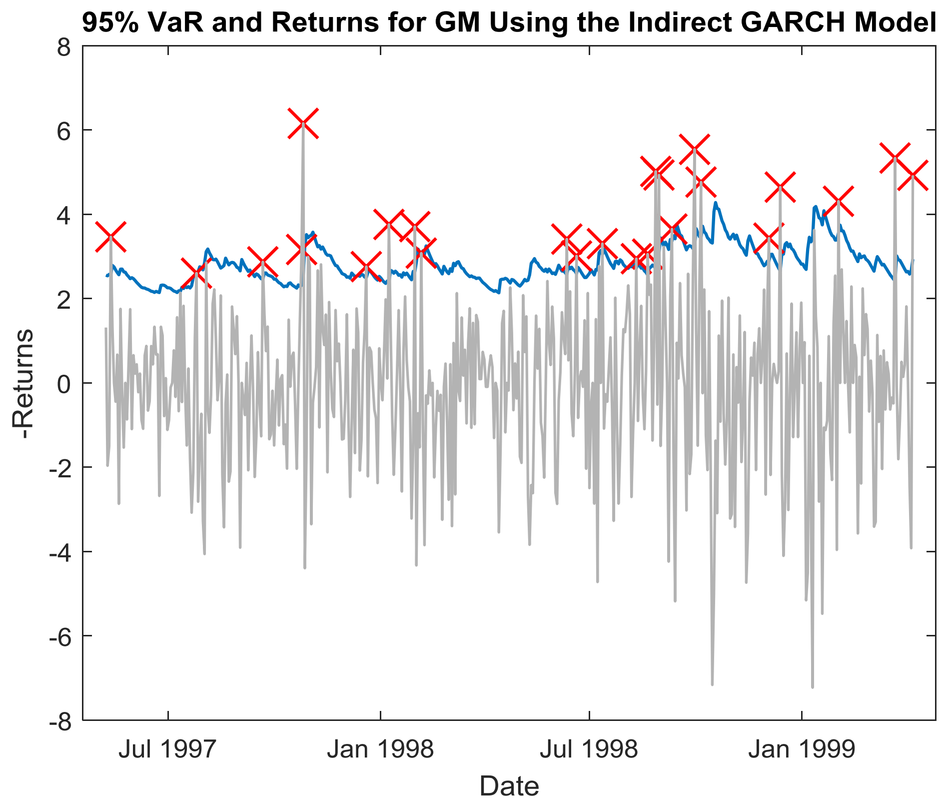


Table 6 *Results of the model suitability test statistics for the 99% VaR models when used on the data from 1986 – 1999*

		1986 - 1999 Data				
		Hits OOS (%)	Coverage Test (*p* values)	Accept at 5% CL?	DQ Test (*p* values)	Accept at 5% CL?
Symmetric Absolute Value	GM	1.0	1.0000	✓	0.7998	✓
	IBM	1.6	0.2149	✓	0.0350	✗
	S&P 500	1.0	1.0000	✓	0.0029	✗
Asymmetric Slope	GM	1.4	0.3966	✓	0.9472	✓
	IBM	1.6	0.2149	✓	0.0427	✗
	S&P 500	1.6	0.2149	✓	0.0476	✗
Indirect GARCH	GM	1.2	0.6630	✓	0.9334	✓
	IBM	1.6	0.2149	✓	0.0349	✗
	S&P 500	1.8	0.1060	✓	0.0309	✗
Adaptive	GM	1.8	0.1060	✓	0.0017	✗
	IBM	1.8	0.1060	✓	0.0020	✗
	S&P 500	1.2	0.6630	✓	0.0035	✗

These results indicate that all the 99% VaR models produced a statistically appropriate number of exceedances in the out-of-sample data. However, very few models were able to pass the DQ test at the 5% significance level. Engle and Manganelli (2004) suggested that perhaps the *p* values could be tested at the 1% level for the DQ test. This would mean that the model would be rejected only if the *p* value is less than 0.01. If we were to relax the significance level of the DQ test here, 8 of the 12 trials would meet the suitability requirements. The Adaptive Model would still fail in all cases, indicating that perhaps it does a poor job of describing the tail dynamics at the 99% VaR level (Engle & Manganelli, 2004). As in the 95% VaR trials, the first 3 models performed best when applied to GM, as indicated by the high *p* values for both tests. Figure 7 shows the 99% VaR as calculated by the Adaptive model and actual returns for GM. Note how the number of exceedances is appropriate, but they again exhibit clustering, particularly shortly after July 1998. This may explain why the model would pass the Coverage Test and fail the DQ test. For comparison, Figure 8 shows the 99% VaR and returns for GM, as calculated by the Asymmetric Slope model. Note how the number of exceedances is appropriate, but they are less clustered than in Figure 7. This may explain why the model would pass both the Coverage and DQ tests.

Figure 7 *99% VaR and Returns for GM when the Adaptive Model is applied.*

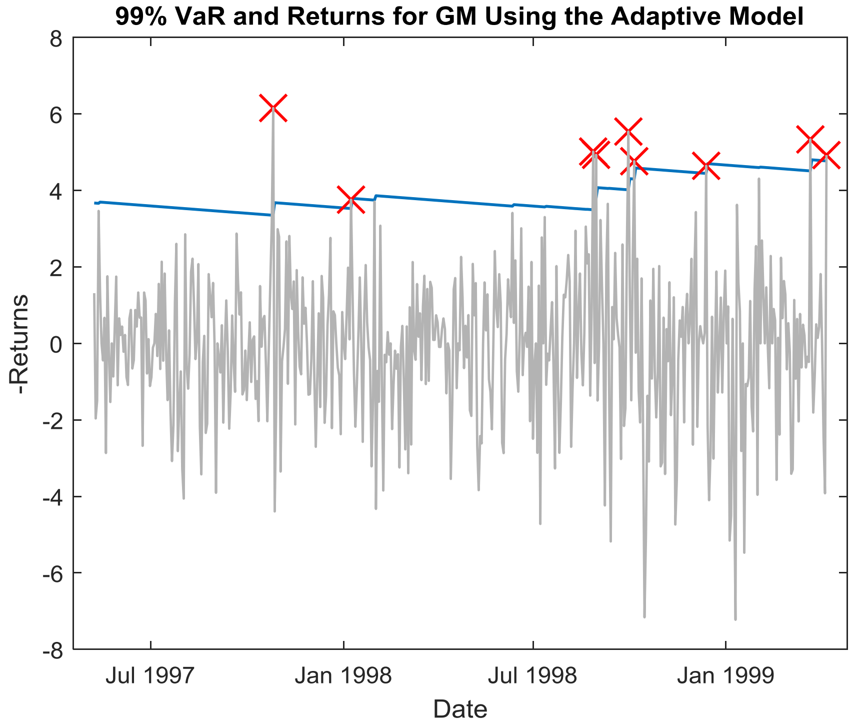


Figure 8 *99% VaR and Returns for GM when the Asymmetric Slope Model is applied.*

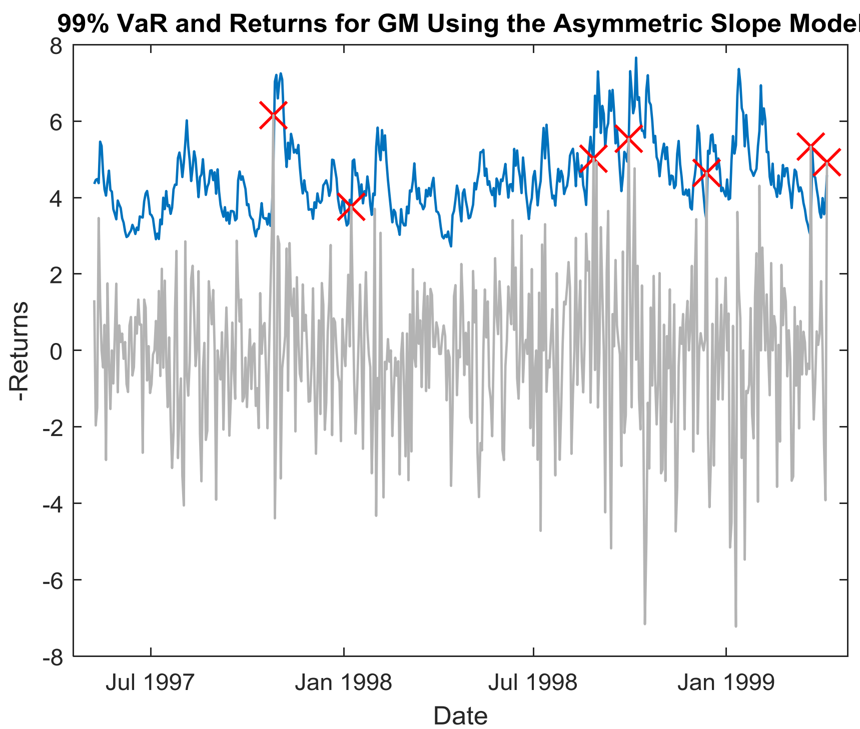


Table 7 contains the data for the test statistics for the 95% VaR models when applied to the out-of-sample data from 1996 – 2009, which includes the recent financial crisis.

Table 7 *Results of the model suitability test statistics for the 95% VaR models when applied to the data from 1996 – 2009*

		1996 - 2009 Data				
		Hits OOS (%)	Coverage Test (*p* values)	Accept at 5% CL?	DQ Test (*p* values)	Accept at 5% CL?
Symmetric Absolute Value	GM	11.0	0.0000	✗	0.0000	✗
	IBM	6.8	0.0792	✓	0.0274	✗
	S&P 500	10.4	0.0000	✗	0.0000	✗
Asymmetric Slope	GM	10.0	0.0000	✗	0.0001	✗
	IBM	6.0	0.3192	✓	0.0069	✗
	S&P 500	9.8	0.0000	✗	0.0000	✗
Indirect GARCH	GM	9.2	0.0001	✗	0.0016	✗
	IBM	10.0	0.0000	✗	0.0000	✗
	S&P 500	8.8	0.0004	✗	0.0012	✗
Adaptive	GM	8.8	0.0004	✗	0.0001	✗
	IBM	7.6	0.0129	✗	0.0392	✗
	S&P 500	8.0	0.0045	✗	0.0237	✗

The results show that nearly all the models fail the coverage test. Only IBM is able to pass the coverage test at the 95% level for the Symmetric Absolute Value and Asymmetric Slope models. The results also show that some of the models have a VaR exceedance percentage that is higher than 10%, when we expect it to be 5% for these 95% VaR models. As shown in Figures 9 and 10, IBM was exposed to the dot-com crash which occurred around 1999-2001, when it exhibited high volatility. This high volatility would have been included in the model estimation. Conversely, GM and the S&P 500 Index were not exposed to the same levels of volatility. This may explain why the models were able to pass the Coverage test for IBM at the 95% significance level.

Figure 9 *Stock Price Chart of IBM.*

Note the high price volatility from 1999 – 2001.

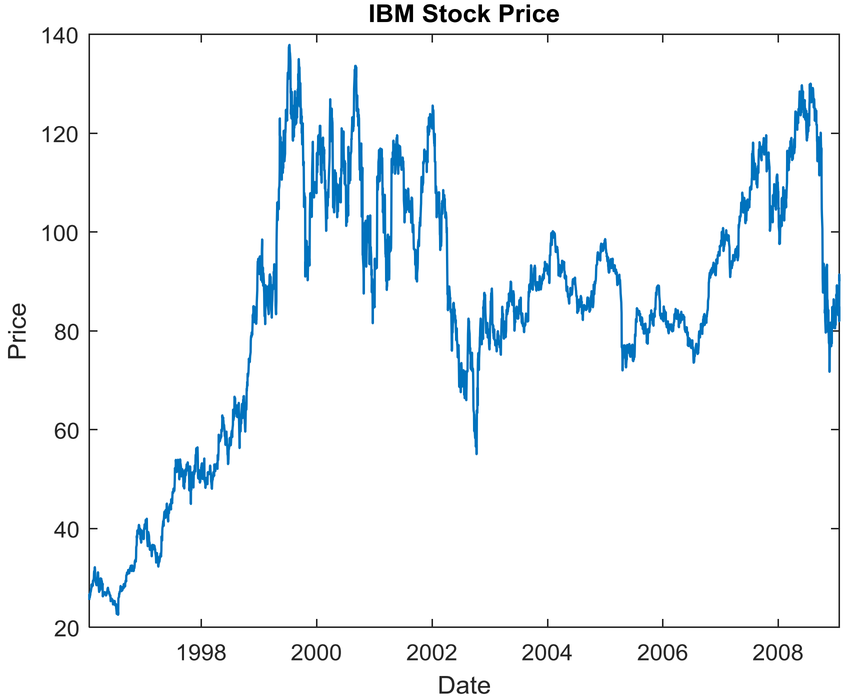
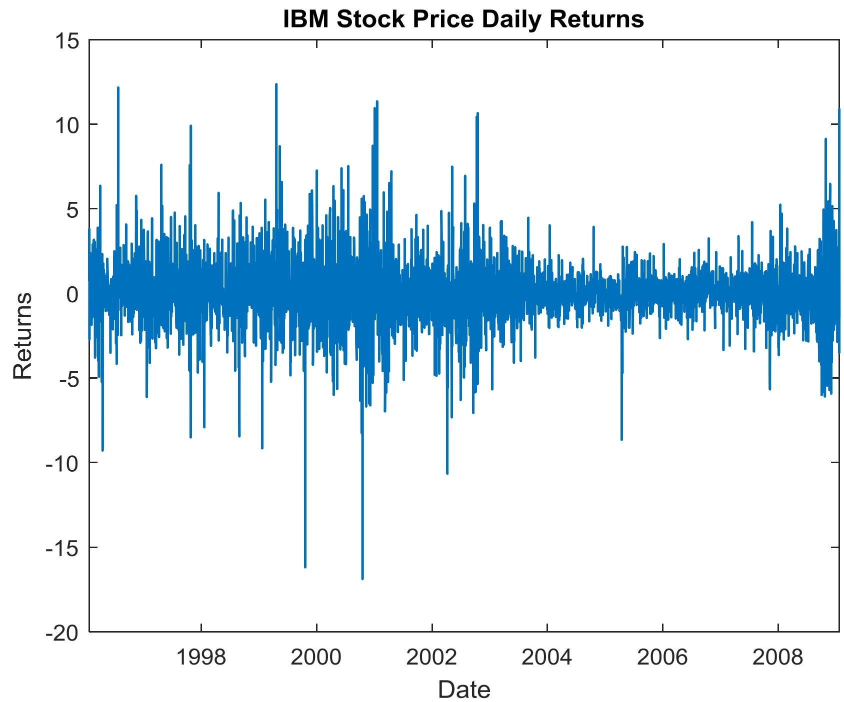


Figure 10 *Daily Returns of IBM Stock*



Additionally, none of the models were able to pass the DQ test at the 95% confidence level. The models likely failed in IBM's case due to exceedance clustering, while the other models failed due to a combination of inadequate coverage and exceedance clustering. Figure 11 shows the 95% VaR and actual returns for IBM using the Asymmetric Slope model. Note that despite producing an appropriate number of exceedances, the exceedances are heavily clustered, especially from Aug 2008 to Sep 2008.

Figure 11 95% VaR and Returns for IBM when the Asymmetric Slope Model is applied.

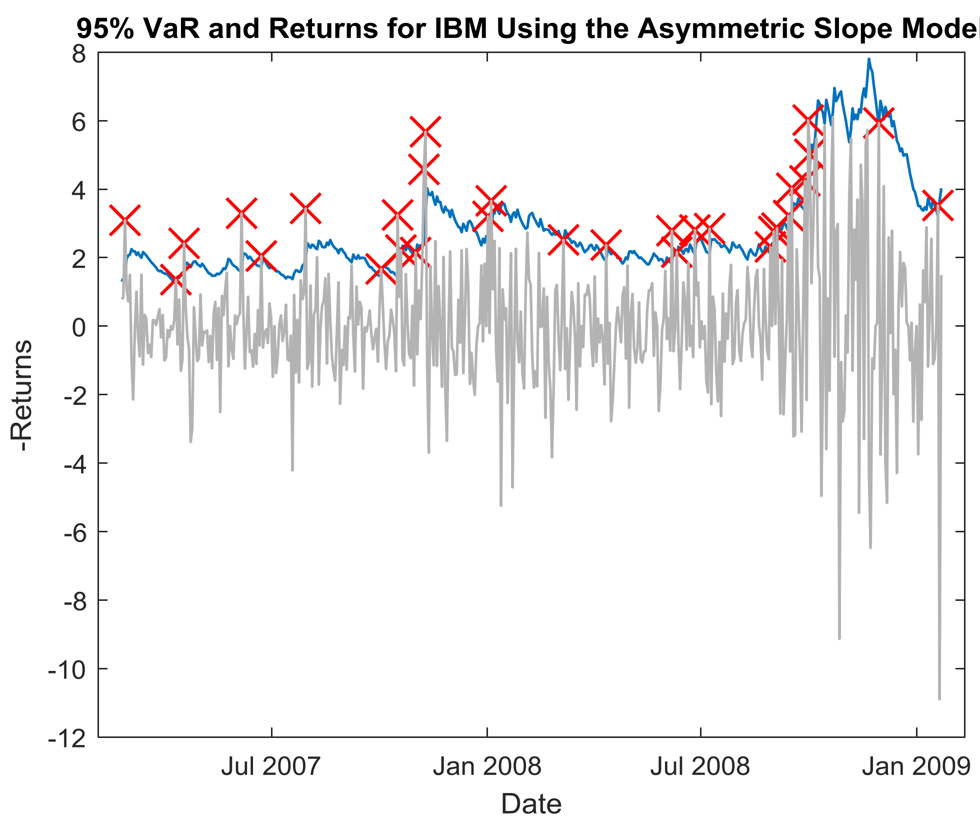
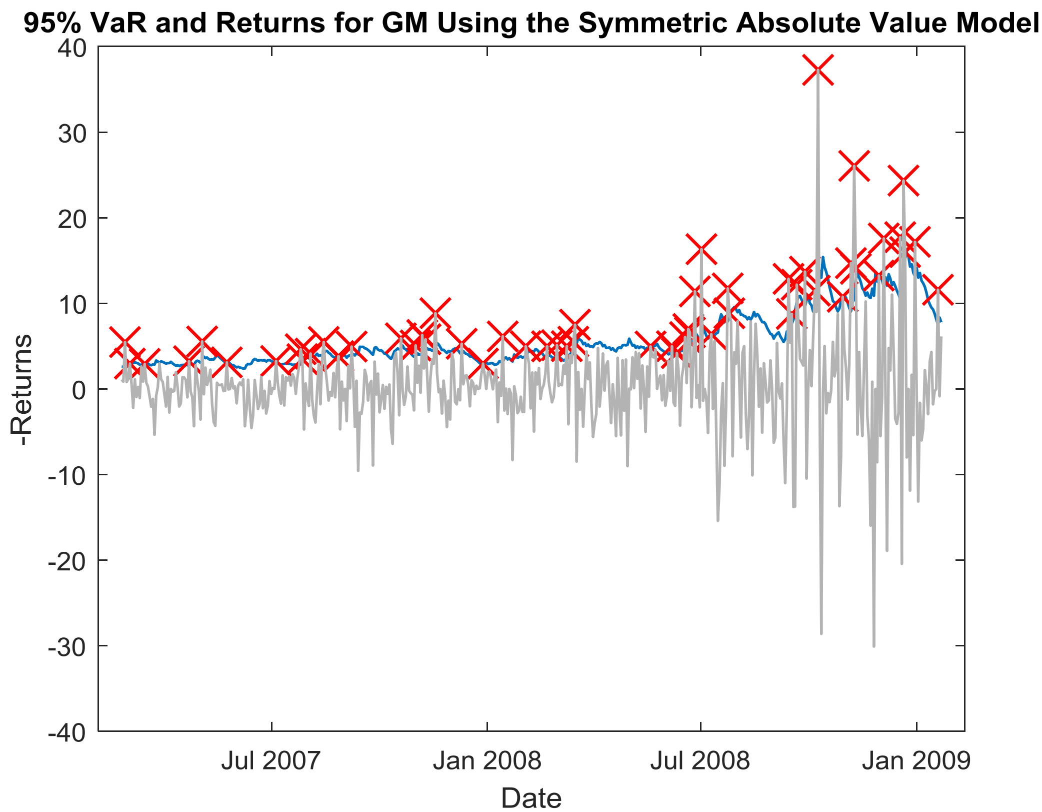


Figure 12 shows the 95% VaR and actual returns for GM using the Symmetric Absolute Value model.

Figure 12 *95% VaR and Returns for GM when the Symmetric Absolute Value Model is applied.*



In Figure 12, note how in addition to the model failing to generate an appropriate number of VaR exceedances, the exceedances it does produce are heavily clustered throughout the time series, indicating poor model performance. This is a stark contrast to the 1986 – 1999 results, where this model exhibited excellent performance when applied to GM. Clearly, these models do not perform well at the 95% VaR level when market conditions are changing.

Table 8 contains the data for the test statistics for the 99% VaR models when applied to the out-of-sample data from 1996 – 2009, which includes the recent financial crisis.

Table 8 *Results of the model suitability test statistics for the 99% VaR models when applied to the data from 1996 – 2009*

		1996 - 2009 Data				
		Hits OOS (%)	Coverage Test (*p* values)	Accept at 5% CL?	DQ Test (*p* values)	Accept at 5% CL?
Symmetric Absolute Value	GM	4.8	0.0000	✗	0.0000	✗
	IBM	1.0	1.0000	✓	0.0002	✗
	S&P 500	4.2	0.0000	✗	0.0000	✗
Asymmetric Slope	GM	3.8	0.0000	✗	0.0000	✗
	IBM	1.4	0.3966	✓	0.6464	✓
	S&P 500	3.8	0.0000	✗	0.0000	✗
Indirect GARCH	GM	2.2	0.0199	✗	0.0353	✗
	IBM	1.4	0.3966	✓	0.0092	✗
	S&P 500	3.8	0.0000	✗	0.0000	✗
Adaptive	GM	3.4	0.0000	✗	0.0000	✗
	IBM	1.6	0.2149	✓	0.0151	✗
	S&P 500	2.6	0.0027	✗	0.0000	✗

Except for IBM, the results show that all of the 99% VaR models fail the coverage test at the 95% significance level. The results also show that the models have VaR exceedance percentages ranging from higher than 2% to close to 5%, when we expect it to be 1% for each of these models. The models were likely able to pass the Coverage test when applied to IBM because of the high volatility it experienced around 1999 – 2001, as shown previously in Figures 9 and 10.

Figure 13 shows the 99% VaR and actual returns for IBM using the Asymmetric Slope model. Note how the exceedances exhibit little clustering and there are no VaR exceedances after July 2008, as the model rapidly adapted to the changing market conditions.

Figure 14 shows the 99% VaR and actual returns for the S&P 500 Index using the Asymmetric Slope model.

Figure 13 99% VaR and Returns for IBM when the Asymmetric Slope Model is applied.

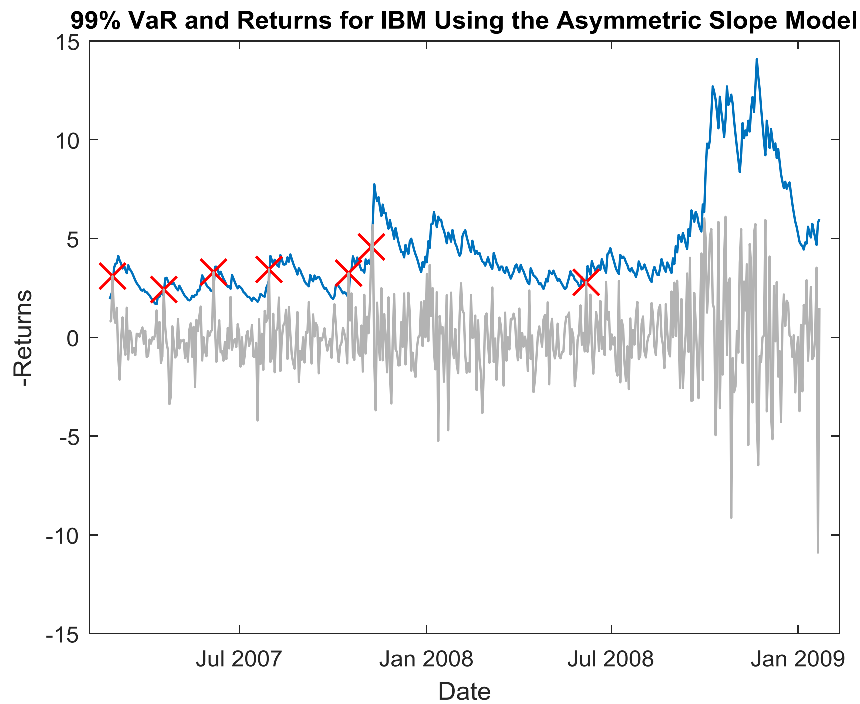
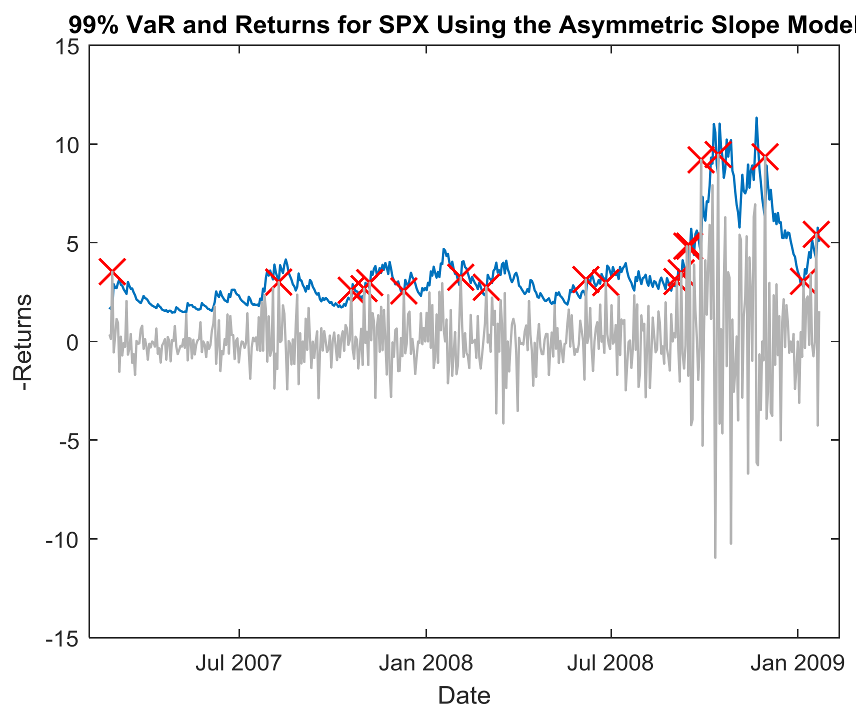


Figure 14 99% VaR and Returns for the S&P 500 Index when the Asymmetric Slope Model is applied.



Notice in Figure 14 how we would expect to see 5 exceedances in a 99% VaR model applied to 500 data points, but we actually observe 19 exceedances. Additionally, these exceedances exhibit significant clustering. This may explain why this model failed the coverage and DQ tests when it was applied to the S&P 500 Index.

4.3 Discussion of Results

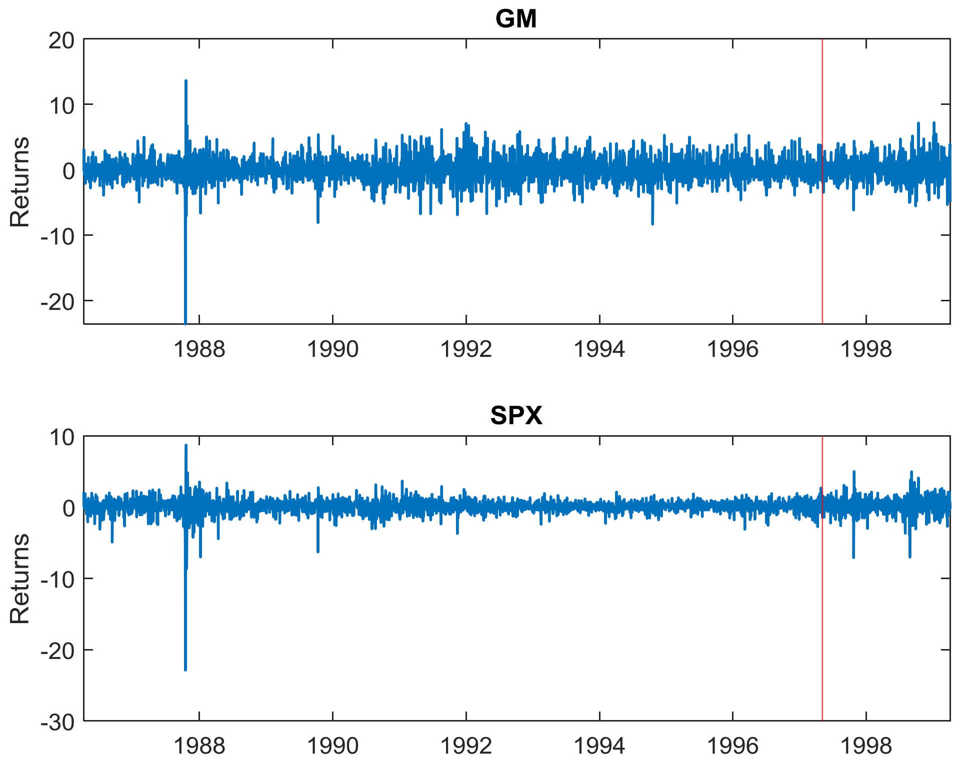
While the models performed well when the historical in-sample time period was representative of the market conditions in the out-of-sample time period, significant differences arise when the market conditions are different.

These results show how when the market conditions are changing from stable conditions to more stressed conditions, risk measurement methods such as VaR that rely heavily on historical data tend to perform very poorly if the conditions represented by the historical data differ significantly from the new conditions. What is important is that there were no warning signs when estimating the models for these new conditions. The results of the model estimation still indicated that the model parameters were statistically significant, and the optimization routine generated an appropriate number of in-sample VaR exceedances. The poor model performance was not observed until the models were applied to the stressed market conditions, when they failed to appropriately capture the market risks and generated far more VaR exceedances than expected. If these models were used to determine a financial institution's capital reserves, there would not be sufficient capital available during the stressed market conditions, because losses that exceed the VaR level would have been far more frequent than expected. The results show that these losses could have occurred anywhere from two to five times as frequently as expected at the 99% VaR level. This could lead to major liquidity problems and systemic crises, as the financial institution scrambles to recover.

In these tests, the models may have been able to perform better at the 99% VaR level for IBM during the crisis because of IBM's prior exposure to stressed market conditions several years earlier during the dot-com crash. The models may have performed poorly for GM and the S&P 500 Index during the stressed period because the historical time period for them did not include a sufficiently stressed period. This meant that the models would not have been calibrated appropriately for stressed conditions, and thus failed to adequately perform under those conditions.

When looking at the 99% VaR for the 1986 – 1999 data, we observe that the models performed well for GM, but not so well for IBM or the S&P 500. This may have been because the out-of-sample volatility for GM was similar to its in-sample volatility, while the out-of-sample volatility for the other securities was higher than their in-sample volatility. This is represented in Figure 15, with the vertical red line separating the in-sample data from the out-of-sample data. Additionally, note how even though return data from the stock market crash of 1987 was included as part of the in-sample data, this was insufficient to calibrate the model appropriately for the S&P 500 Index. This may have been because the contribution from the long period of low volatility from 1990 – 1999 diminished the contribution from the crash in 1987.

Figure 15   Returns Comparison for GM and the S&P 500 Index



The DQ test used in this research tested for serial correlation in the previous four lagged exceedances, and subjected any detected autocorrelation to the same penalty, regardless if it happened one time step earlier, or four time steps earlier. It is unclear if this is the best way to evaluate this condition. It may be that the models could be better evaluated by looking at fewer (or more) lagged values, or maybe the penalty for autocorrelation could be different depending on how recently it was detected.

Another potential concern is related to the inclusion of returns on holidays in the time series. Returns on these days is always 0%, and inclusion of these data points may adversely impact the model's accuracy because these values could alter the distribution of returns. With around 10 holidays in every year, the testing period of 14 years could include returns for 140 holidays. With 252 trading days in a year, these holidays contribute more than half a year of predictable 0% returns to the full data set. Thus, it may be more suitable to remove these entries from the time series data. However, in the interest of proper scientific experimentation where as many variables are held fixed as possible, this research included holidays in the time series data in order to be consistent with the original research.

Using the code and data provided by the authors of the original paper, it was possible to replicate exactly the published results for the Asymmetric Value, Indirect GARCH, and Adaptive models. However, the results of the Symmetric Absolute Value model could not be replicated exactly, although this did not result in changes to the model's suitability. E-mail discussions with the authors suggested that the discrepancies were likely due to MATLAB's recent changes to some of its optimization algorithms (S. Manganelli, personal communication, May 22, 2018).

5. Conclusion

Despite its well-known shortcomings and limitations, Value-at-Risk continues to be used as a common measure of market risk. This research examined four VaR models that were previously published in the academic literature. The models were applied to the same data in the published research as well as data from a more recent time period. Results from both data sets were compared, and the models were critically evaluated for their suitability under stressed market conditions. It was found that the historical data used in the estimation of a VaR model significantly impacts the model's ability to capture market risks. Specifically, when the market conditions change significantly to stressed conditions, the VaR models fail to adequately capture the market risks. This could be extremely important if the VaR measures are used to determine the riskiness of an investment, or to calculate a financial institution's capital reserves, because the failure to adequately measure the risks of an investment could expose a financial institution to serious problems. This is one reason why regulators and banking supervisors are encouraging banks to use techniques such as stress testing when calculating their capital reserve requirements (BIS, 2009).

Reference List

Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent measures of risk. *Mathematical finance, 9*(3), 203-228.

BIS. (1995). *An Internal Model-Based Approach to Market Risk Capital Requirements.* Basel: Bank for International Settlements.

BIS. (2009). *Principles for sound stress testing practices and supervision.* Basel: Bank for International Settlements.

Bloomberg L.P. (2018). *Stock Price data for MTLQQ, IBM, and SPX 4/7/1986 to 1/22/2009.* Retrieved from Bloomberg Database.

Daníelsson, J. (2011). *Financial risk forecasting: The Theory and Practice of Forecasting Market Risk, with Implementation in R and MATLAB.* Hoboken: John Wiley & Sons Ltd.

Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business & Economic Statistics, 22*(4), 367-381.

Huang, D., Yu, B., Fabozzi, F. J., & Fukushima, M. (2009). CAViaR-based forecast for oil price risk. *Energy Economics, 31*(4), 511-518.

Jeon, J., & Taylor, J. W. (2013). Using CAViaR Models with Implied Volatility for Value-at-Risk Estimation. *Journal of Forecasting, 32*(1), 62-74.

Koenker, R., & Bassett Jr, G. (1978). Regression Quantiles. *Econometrica, 46*(1), 33-50.

Kuester, K., Mittnik, S., & Paolella, M. S. (2006). Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics, 4*(1), 53-89.

Manganelli, S. (2018, November 26). *Simone Manganelli | Research*. Retrieved from Simone Manganelli: http://www.simonemanganelli.org/Simone/Research.html

Taylor, J. W. (2008). Estimating Value at Risk and Expected Shortfall Using Expectiles. *Journal of Financial Econometrics, 6*(2), 231-252.