STAT 461 Project
GARCH Forcasts of S&P 500 and VIX Volatility

Shaun Campbell
scampbell22@wisc.edu

December 11, 2019

1 Executive Summary

Generalized autoregressive conditional heteroskedasticity (GARCH) models are used in finance to forecast volatility. In this report, GARCH models are created to forecast the volatility of the S&P 500 and the VIX, which are two of the most-watched market indexes. Being able to forecast the volatility of the market has definitive implications in terms of risk management. However, this report attempts to develop a trading strategy based on the volatility forecasts and the relationship between volatility and returns.

Daily data is retrieved for the VIX, S&P 500, the VXX (a VIX ETF) and the SPY (an S&P ETF) going back to 2010. The data is split into training, validation, and test sets. Preliminary analysis is performed to ensure GARCH models are appropriate for the time series. Then GARCH models are fitted to the training data for the S&P 500 and VIX and model diagnostics are performed to ensure the models are adequate.

Once the models are developed, a trading strategy is designed around the volatility forecasts. The strategy takes advantage of the "leverage effect", i.e. the negative relationship between volatility and returns. Since the S&P and VIX tend to trend up and down respectively over time, positions counter to these trends are not taken. Instead, long and short positions are respectively taken in the S&P and VIX only when volatility is projected to decrease. This should allow market-friendly positions during times of low volatility — implying steady positive returns in the S&P and negative returns in the VIX — and no positions during periods of high volatility to avoid losses. The objective is to capture positive gains in the positions while minimizing drawdowns.

The trading strategy, once tested, has mixed results. The strategy performs very well on the validation set which saw very low volatility. Performance was not as good on the test set, which saw heavily increased volatility in comparison to the validation set. This suggests, among other things, that the model performs best during periods of low volatility. Overall, the strategy does feature lower drawdowns than most benchmarks and is concluded to have potential but in need of further testing going forward. There are also simpler strategies that may achieve similar results.

The analysis set forth in this report shows that in addition to risk management, GARCH can have an important role in trading. Traders could adjust their positions based on GARCH and use volatility forecasts to choose entry points. These implications confirm the versatility and power of GARCH models.

2 Introduction

There are many applications of GARCH, primarily in risk management. Though asset returns are very difficult to forecast, volatility can be forecast because it is usually partially dependent on past volatility. Volatility has a tendency to cluster and features exponential decay following a spike.

Though the ability to forecast volatility does not have any immediate implications for trading strategies, one may develop one by utilizing the relationship between volatility and returns. The so-called "leverage effect" is empirical evidence that asset returns are negatively correlated with their volatility. If this relationship is strong enough, volatility predictions could potentially be used as a proxy for return predictions. The purpose of this report is to investigate this topic by developing GARCH models for two of the most-watched indexes — the VIX and the S&P 500 — then implementing a trading strategy using the models.

3 Data Description

Data was downloaded from Yahoo Finance using the R package quantmod for indexes and ETFs for the VIX and S&P 500 from January 2010 to November 2019. The data includes open, high, low, close, adjusted close, and volume data for the indexes/ETFs. The S&P 500 ETF is SPY and the VIX ETF is VXX which holds a long position in 1 month VIX futures.

Log-returns were calculated for each index/ETF and the log-returns were used for analysis. For analysis of volatility, σ_t values are approximated with $|r_t|$.

The data is split into train, validation, and test sets. The training set includes data from 2010-01-04 to 2016-11-28, the validation set includes data from 2016-11-28 to 2017-11-28, and the test set includes data from 2017-11-28 to 2019-11-28.

4 Preliminary Analysis

The first objective is to develop a GARCH model to predict future volatility of the S&P 500 and the VIX indexes. First, preliminary analysis of the return series for the indexes is conducted. Plots to visualize the return series are created for the S&P 500 and the VIX and are shown in Figure 1.

Next, to look for autocorrelation in the return series, ACF and PACF plots are created and are shown in Figure 2. The ACF plots show no significant autocorrelation for either index. However, both PACF functions show significant autocorrelation. Due to this, an ARIMA model could be fit to the data to potentially forecast returns of the series. However, since this report is focusing on GARCH models, the ARIMA models will not be fit.

ACF and PACF are then constructed for the squared return series to investigate the autocorrelation of the volatility of the indexes (shown in Figure 3). The ACF and PACF both show significant autocorrelation in the data for both indexes. This signifies that a GARCH model will be appropriate to forecast future volatility.

5 GARCH Models

5.1 S&P 500 GARCH Model Fitting

To begin fitting a GARCH model for the S&P 500, some reasonable possible GARCH models are identified. Since GARCH(1,1) models are the most commonly used and usually sufficient, the order of the GARCH model will be set to (1,1). The distributions that will be investigated are normal,

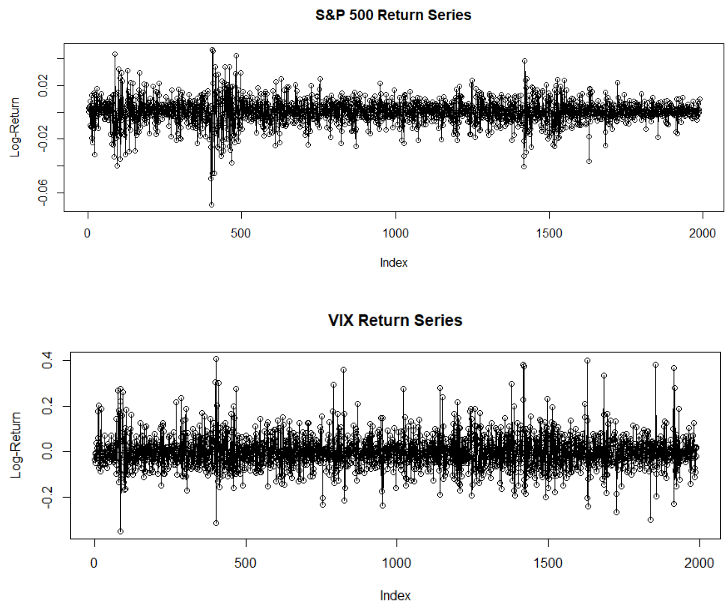

Figure 1: S&P 500 and VIX Return Series

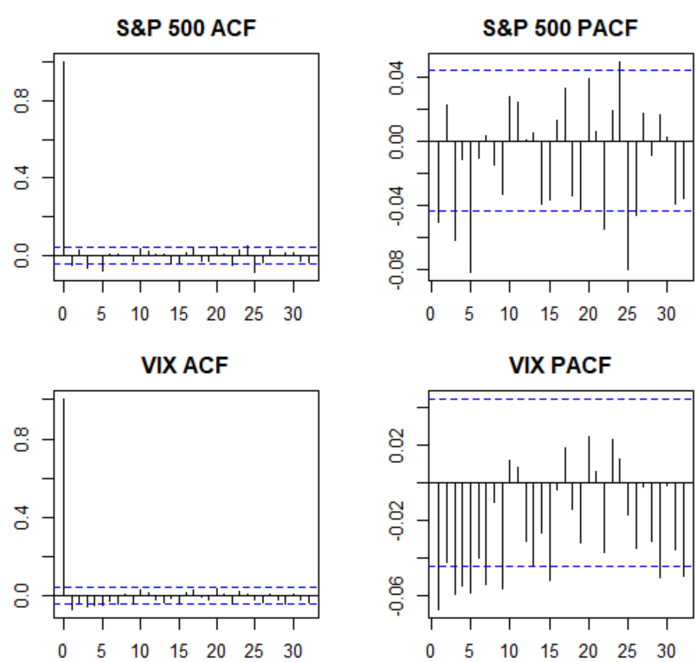

Figure 2: Autocorrelation and Partial-Autocorrelation Functions for S&P 500 and VIX

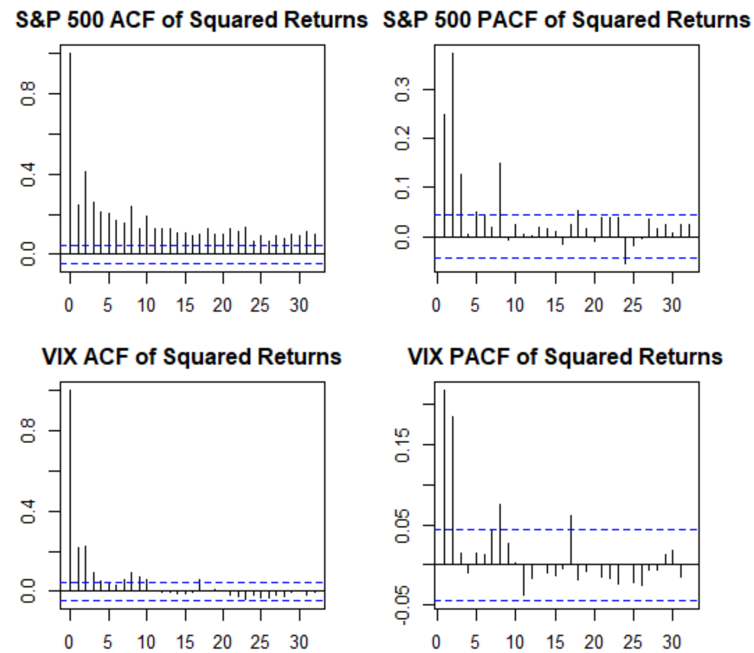

Figure 3: Autocorrelation and Partial-Autocorrelation Functions of Squared Returns for S&P 500 and VIX

Student t, and skewed Student t. Finally. normal GARCH, exponential GARCH (EGARCH) and threshold GARCH (TGARCH) will be explored, since the latter two are designed to account for the leverage effect.

To decide upon the distribution, the skewness and kurtosis of the volatility of the training data returns is calculated. $|r|$ is used as a proxy for volatility. The distribution of the volatility can be seen in Figure 4 and the statistics for the skewness and kurtosis are displayed in Table 1. As shown by the density curve and confirmed by the statistics, the skewness and excess kurtosis of the volatility are both high. The skewness is positive meaning the volatility is right-skewed. Because of this, a normal distribution would not be appropriate. A Student t distribution would account for the kurtosis, but a skewed Student t distribution would account for both skewness and kurtosis. Therefore, a skewed Student t distribution is chosen.

Measure	Value
Skewness	2.386
Excess Kurtosis	6.250

Table 1: S&P 500 σ Skewness and Kurtosis

Next, possible GARCH models are explored. GARCH, EGARCH, and TGARCH are fit to the training data and then the information criteria statistics are calculated. The AIC and BIC values for each model are shown in Table 2. The TGARCH model minimizes both the AIC and BIC so it is chosen for the GARCH model. A (1,1) TGARCH model with a skew Student t distribution is fit to the training data. The resulting equation and estimated coefficients of the TGARCH model

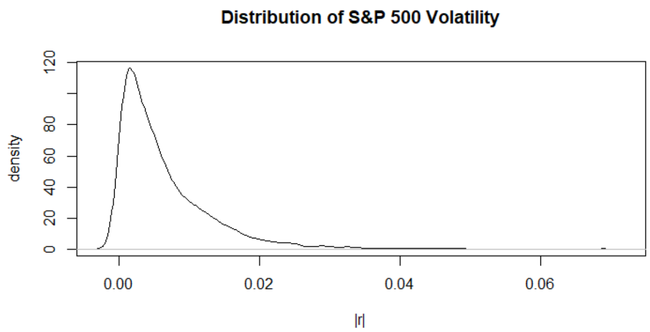

Figure 4: S&P 500 σ Density Curve

Order		GARCH	EGARCH	TGARCH
(1,1)	AIC	-6.718183	-6.782772	-6.788374
	BIC	-6.674198	-6.735644	-6.741247

Table 2: S&P 500 Choice of GARCH Model

are shown below and in Table 3.

$$\sigma_t^2 = \alpha_0 + (\alpha_1 + \gamma_1 N_{t-1})a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

	Estimate	Std. Error	p-value
α_0	0.000514	0.000423	0.224
α_1	0.132324	0.012902	0.000
β_1	0.847262	0.042869	0.000
γ_1	1.000000	0.050256	0.000
skew	0.821245	0.066318	0.000
shape	6.704627	1.700647	0.000081

Table 3: Coefficients for TGARCH(1,1) Model for S&P 500

To visualize the fit of the model, the model is used to produce 1-step ahead forecasts of volatility on the validation set. Then the forecasts are plotted with the realized values of volatility ($|r_t|$). All coefficients are statistically significant except for the intercept term α_0, and noteably the model appears to generally overestimate the mean of the volatility. This will be explored further in future sections. The predicted volatility values still follow the same general trend as the realized volatility. The plot can be seen in Figure 5.

5.2 S&P 500 GARCH Model Diagnostics

Several plots are generated to check model assumptions for the S&P 500 TGARCH model and perform model diagnostics. The diagnostics plots are shown in Figure 6. The residuals of the

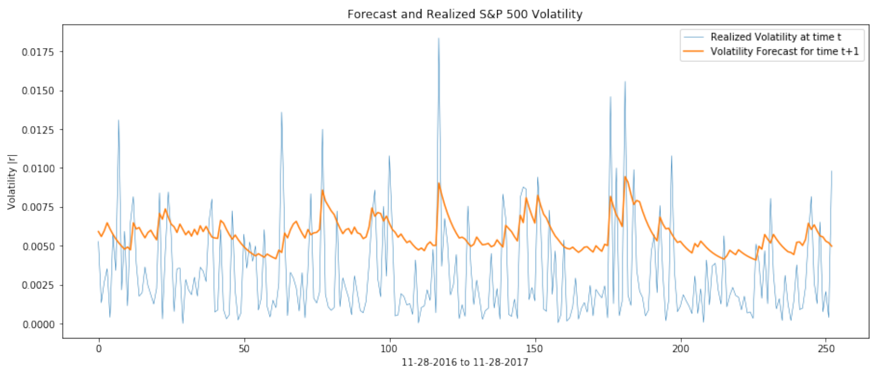

Figure 5: S&P 500 Realized Volatility and TGARCH(1,1) Volatility Forecast

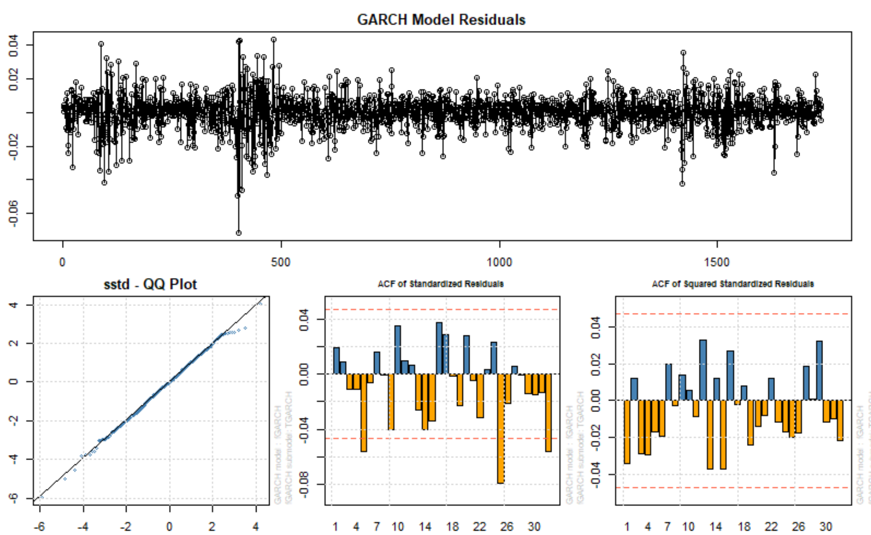

Figure 6: Diagnostics Plots for S&P 500 TGARCH Model

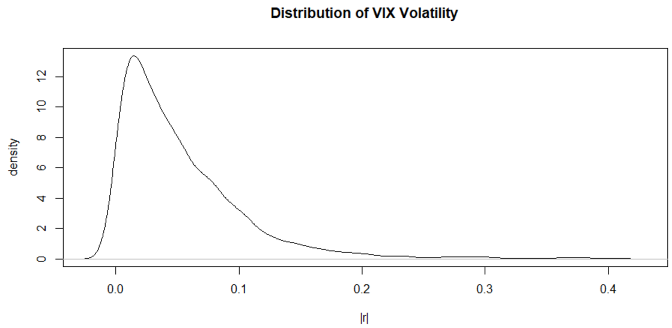

Figure 7: VIX σ Density Curve

model appear to have constant variance, so homoscedasticity of the residuals is satisfied. The QQ-Plot of standardized residuals is also satisfactory. The linear pattern of the residuals along the QQ line implies that a skewed Student t distribution is a proper choice.

There appears to be some autocorrelation in the residuals in the ACF plot. This is confirmed by an Ljung-Box test on the residuals which finds that some lags have significant autocorrelation in the residual series. However, there is no clear autocorrelation in the ACF plot of squared residuals, which is also confirmed with an Ljung-Box test. The autocorrelation of the residuals occur at lag-5, 25, and 32. Lag-25 and 32 are very high order lags and can be dismissed as occurring randomly. Lag-5 is also a fairly high order lag, and it is the only significant autocorrelation. Since no squared residuals have significant lags, the model is concluded to be satisfactory.

5.3 VIX GARCH Model Fitting

The same method is used to fit a GARCH model for the VIX. First the skewness and kurtosis are calculated for the volatility of the VIX returns. The values are similar to those for the S&P 500 and are displayed in Table 4. The density plot (shown in Figure 7) and skewness value confirm a positive right skew and the kurtosis is quite large. Therefore a skewed Student t distribution will be used.

Measure	Value
Skewness	2.344
Excess Kurtosis	5.487

Table 4: VIX σ Skewness and Kurtosis

Again, (1,1) GARCH, EGARCH, and TGARCH models are explored to find the best model to

Order		GARCH	EGARCH	TGARCH
(1,1)	AIC	-2.597728	-2.648302	-2.640078
	BIC	-2.572582	-2.620012	-2.611788

Table 5: VIX Choice of GARCH Model

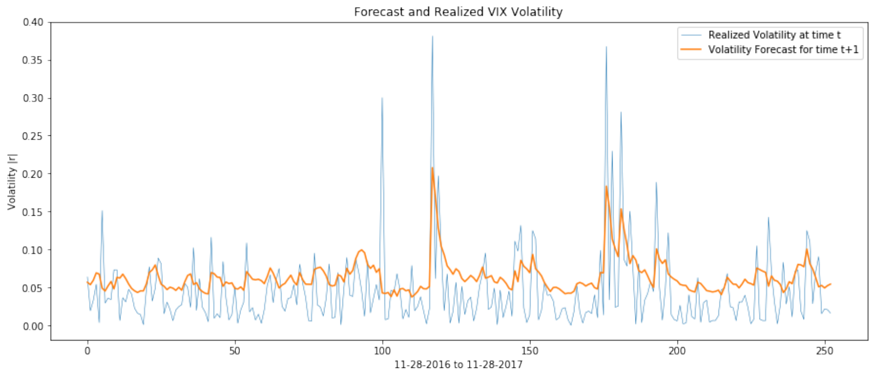

Figure 8: VIX Realized Volatility and EGARCH(1,1) Volatility Forecast

fit the training data. The information criteria statistics are calculated and shown in Table 5. The model that minimizes both the AIC and BIC is the EGARCH model so it is adopted to model the volatility of the VIX. The resulting model coefficients are displayed in Table 6. All of the model estimations are statistically significant. The model is used to produce 1-step ahead forecasts for the validation set. The resulting fit is shown in Figure 8. The fit appears to be very good. The model does a good job approximating the trend of the volatility and does not appear to overestimate the mean of the volatility by a significant amount.

	Estimate	Std. Error	p-value
α_0	-0.673166	0.034455	0.000000
α_1	0.322020	0.009450	0.000000
β_1	0.875073	0.006737	0.000000
γ_1	0.060837	0.026591	0.022145
$skew$	1.414433	0.044826	0.000000
$shape$	5.838004	0.808526	0.000000

Table 6: Coefficients for EGARCH(1,1) Model forVIX

5.4 VIX GARCH Model Diagnostics

Plots are generated for the VIX EGARCH model to check assumptions and perform model diagnostics. The plots are displayed in Figure 9. Homoscedasticity of residuals is satisfied because the model residuals appear to have constant variance. The QQ-Plot of standardized residuals is also satisfactory and the tight pattern along the QQ line confirms that a skewed Student t distribution is an adequate choice.

The ACF plot shows some autocorrelation in the residuals, which is confirmed by an Ljung-Box test on the residuals. The autocorrelation of the residuals occur at lag-1 and 13. The autocorrelation at lag-1 is somewhat concerning as it might imply that the model is not adquate. However, the autocorrelation value is barely past the significance threshold and there is a lack of autocorrelation in the squared residuals. The lack of autocorrelation in squared residuals is confirmed with an

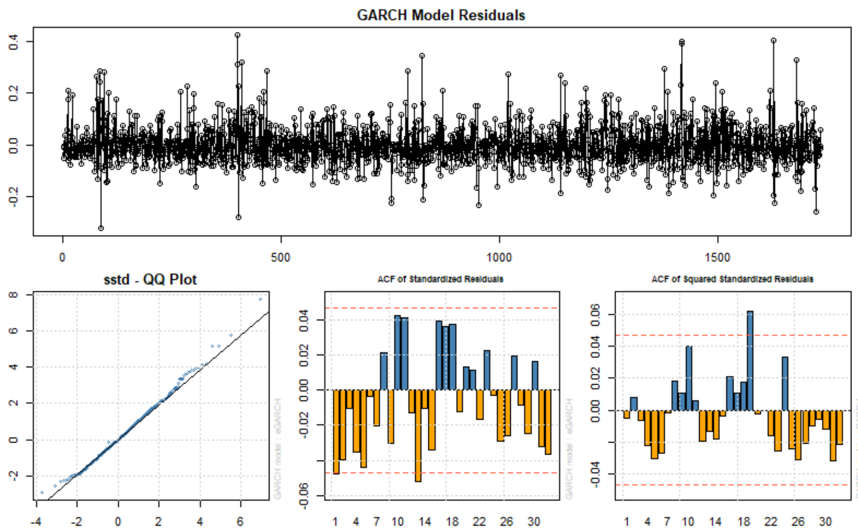

Figure 9: Diagnostics Plots for VIX EGARCH Model

Ljung-Box test. Despite a hint of lag-1 autocorrelation in the residuals, the model is nonetheless concluded to be adequate due to the absence of autocorrelation in the squared residuals.

6 Trading Strategy

To a develop a trading strategy using the GARCH models, the relationships between the volatilities and the returns of the underlying assets are investigated. The correlations between the asset volatilities and returns are calculated and displayed in Table 7. The correlations are fairly weak, but an increase in VIX volatility is associated with a positive change in VIX returns and negative change S&P 500 returns. Likewise, an increase in S&P 500 volatility is associated with a positive change in VIX returns and a negative change in S&P 500 returns. This suggests that if an increase in volatility is predicted, the VIX be longed and the S&P 500 shorted. Conversely, a predicted decline in volatility would suggest a short position in the VIX and a long position in the S&P 500.

Since empirically the VIX trends down and the S&P 500 trends up over time, a short position in the S&P and a long position in the VIX would be risky and one would need to be very confident in a prediction to take these positions. Instead, the predictions will be used to decide when to take short and long positions in the VIX and S&P 500 respectively. Doing so should allow drawdown to be lowered while capturing the upside of the positions.

	σ_{VIX}	σ_{SP500}
r_{VIX}	0.21479	0.13285
r_{SP500}	-0.18363	-0.08041

Table 7: Correlations

The GARCH models are modified slightly to have more predictive utility. Figures 5 and 8 show the GARCH forecasts consistently overestimate the mean of the volatility. If the GARCH models are to be used for making predictions on increases and decreases in volatility, the mean of the forecasts should more accurately reflect the mean of volatility. To account for this, the difference

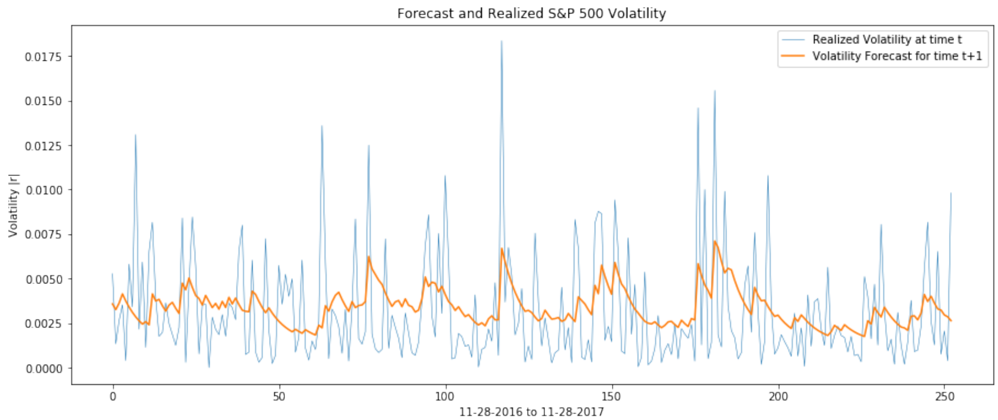

Figure 10: Adjusted GARCH Forecasts for S&P 500

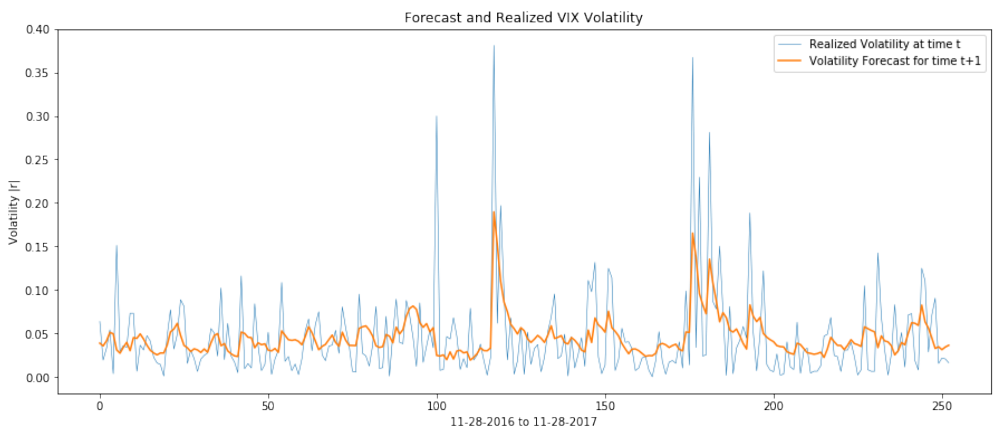

Figure 11: Adjusted GARCH Forecasts for VIX

between mean of the fitted sigma values and the $|r|$ values of the training data is taken. This value is then subtracted from the forecast sigma values in the validation sets. The adjusted fits can be seen in Figures 10 and 11 and appear to be more adequate for predictions.

The prediction rules are designed as follows: if the t+1 forecast volatility at time t is higher than the realized volatility at time t, predict an increase in volatility. If the t+1 forecast volatility at time t is lower than the realized volatility at time t, predict a decrease in volatility. The rules are applied to the validation sets and the results are displayed in Table 8. The results suggest that the model is quite robust at predicting increases and decreases in volatility, with prediction accuracies of 71.43% and 75.79% for the S&P 500 and VIX respectively. The model seems a bit biased towards predicting an increase in volatility. However, the overall accuracy is satisfactory, and if the relationship between volatility and returns can be exploited, then a viable trading strategy should result.

It is unclear whether the S&P 500 GARCH predictions or VIX predictions will be better to generate trading rules. Therefore, 3 different but similar strategies are tested: The S&P 500 GARCH volatility forecasts are used to generate trade signals for both the S&P 500 and VIX [1]; the VIX GARCH volatility forecasts are used to generate signals for both assets [2]; the VIX

S&P 500

	Pred. Inc.	Pred. Dec.
Realized Inc.	112	11
Realized Dec.	61	68

Accuracy: 71.43%

VIX

	Pred. Inc.	Pred. Dec.
Realized Inc.	111	17
Realized Dec.	44	80

Accuracy: 75.79%

Table 8: Confusion Matrices for S&P 500 and VIX Volatility Predictions on Validation Set

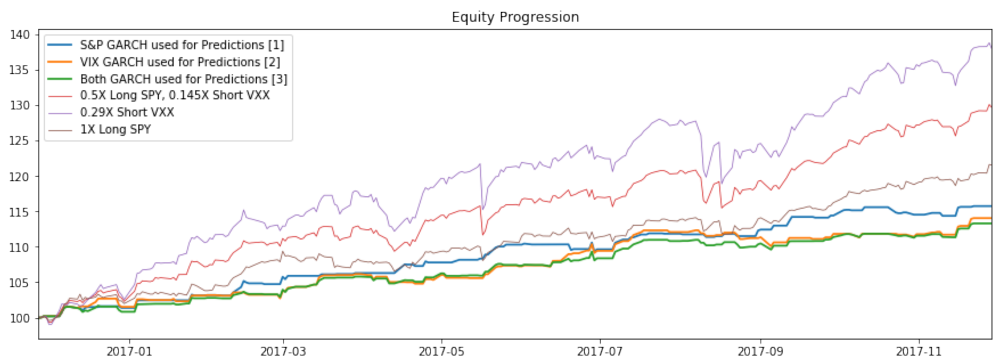

Figure 12: Trade Strategy Equity Curves

GARCH is used to generate signals for the VIX and the S&P 500 GARCH is used to generate signals for the S&P 500 [3]. The VXX ETF is used as a proxy for the VIX and the SPY ETF is used as a proxy for the S&P 500.

For comparison purposes, the trading portfolio and several benchmark portfolios need to be designed. The simple return series from the VXX training data is regressed on market returns to calculate the security's beta. The regression finds a beta of -3.45 for the VXX. For comparison purposes, overall portfolio betas of 1 are desired. Therefore, a weight of $\frac{0.5}{-3.45} = -0.145$ will be assigned to the VXX and a weight of 0.5 assigned to the SPY. This will result in a portfolio beta of $-0.145 * -3.45 + 0.5 * 1 = 1$. To compare with a short VIX position, a weight of $\frac{1}{-3.45} = -0.29$ is assigned to VXX. To compare with a long S&P 500 position, a weight of 1 is assigned to SPY. Finally, a long SPY and short VXX portfolio is created using the same weights as the trading strategy but the trade signals will not be used to enter and exit the positions. All strategies, including the benchmarks, are rebalanced daily.

6.1 Trading Strategy on Validation Set

The trading strategies are tested on the validation data set. The resulting equity curves are shown in Figure 12. It is apparent that the trading strategies do not outperform a long position in

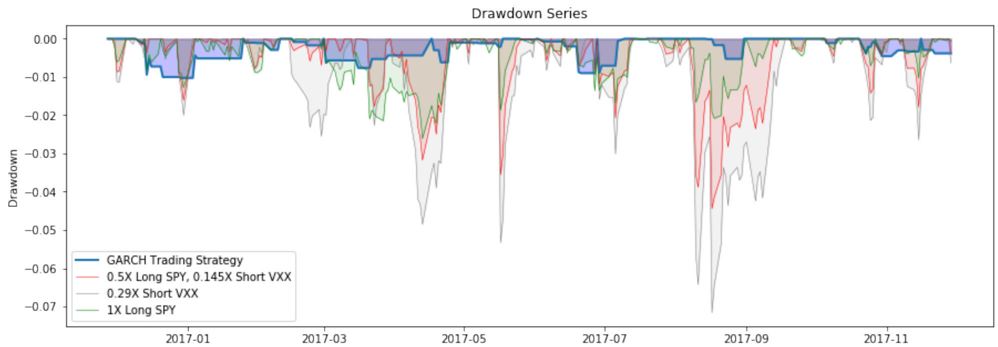

Figure 13: Trade Strategy and Baseline Drawdown Series on Validation Set

SPY in terms of returns. However, that does not mean the strategy is worse. To investigate the performance of the strategy further, several statistics are calculated to compare the strategies.

The S&P GARCH trade signals [1] appear to be the most robust trade signals because the resulting strategy has the highest total return, daily Sharpe ratio, and smallest drawdowns among the three tested strategies. Therefore, this strategy will be used to compare with the benchmark strategies. The trading strategy results in about half the total return as the long SPY/short VXX benchmark, but the annualized daily volatility is much lower (3.95% versus 9.73%). This leads to a higher daily Sharpe in comparison to the benchmarks, including a long SPY position. The trading strategy also has a smaller average drawdown and max drawdown than the long SPY/short VXX benchmark, and even beats the long SPY position in average and max drawdown.

The drawdown series are calculated and plotted for the trading strategy and benchmarks, shown in Figure 13. It is clear that the drawdowns of the trading strategy are much smaller than the benchmarks. Notably, there are several instances in which the benchmarks have very large drawdowns and the trading strategy has not generated a signal so no drawdown occurs. This, along with the statistics, are evidence that the trading strategy achieves its objective of capturing upside in long SPY and short VXX price movements while using volatility forecasts to minimize drawdowns. However, the validation data is only a year worth of data, which is very small. The strategy will be tested on the test set to see if it remains viable.

Stat	S&P Predict [1]	VIX Predict [2]	Both Predict [3]	Long SPY/Short VXX	Short VXX	Long SPY
Total Return	15.76%	14.06%	13.31%	29.66%	37.95%	21.56%
Daily Mean (ann.)	14.60%	13.15%	12.48%	26.26%	32.90%	19.61%
Daily Vol (ann.)	3.95%	4.47%	4.01%	9.73%	13.83%	6.87%
Daily Sharpe	3.69	2.94	3.12	2.70	2.38	2.85
Avg. Drawdown	-0.43%	-0.71%	-0.50%	-0.79%	-1.31%	-0.52%
Max Drawdown	-1.14%	-2.01%	-1.24%	-4.44%	-7.16%	-2.61%

Table 9: Trade Results

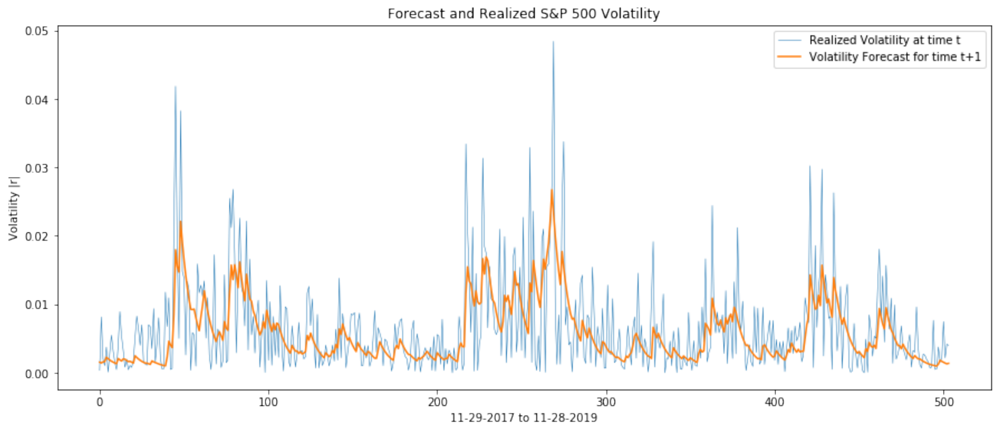

Figure 14: S&P 500 GARCH Forecasts on Test Set

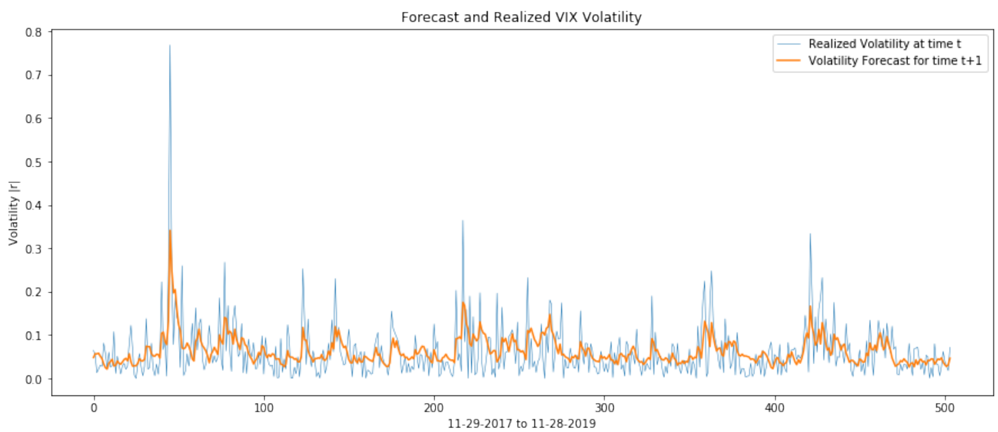

Figure 15: VIX GARCH Forecasts on Test Set

6.2 Trading Strategy on Test Set

The test data set includes data from 2017-11-28 to 2019-11-29. The GARCH models are fit to the test data to visualize the forecasting ability of the models. Again, the mean sigma values for the GARCH forecasts are subtracted from the $|r|$ values from the training and validation sets to shift the GARCH forecasts. The model fits are shown in Figures 14 and 15. The fits appear to be satisfactory and do a good job tracking the volatility trends.

Confusion matrices are generated to see how the GARCH models predict increases and decreases in volatility, shown in Table 10. The models perform very well and score accuracies of 77.14% and 73.36% for the S&P and VIX respectively. These numbers are quite close to the numbers in the validation series.

The trading strategy (using the S&P GARCH to generate trade signals, since it was the best performing method in the validation series) is run on the test data with the same benchmark portfolios. The resulting equity curves are shown in Figure 16. It is immediately apparent that all portfolios experienced more volatility in the test series than in the validation series. In particular, February 2018 featured high levels of volatility and caused big drawdowns. The end of 2018 also

S&P 500

	Pred. Inc.	Pred. Dec.
Realized Inc.	186	63
Realized Dec.	52	202

Test Accuracy: 77.14 %

VIX

	Pred. Inc.	Pred. Dec.
Realized Inc.	204	91
Realized Dec.	43	165

Accuracy: 73.36%

Table 10: Confusion Matrices for S&P 500 and VIX Volatility Predictions on Test Set

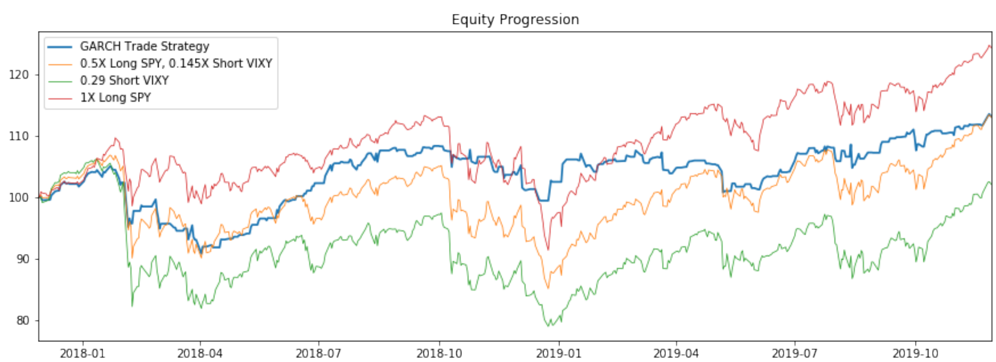

Figure 16: Trade Strategy Equity Curves on Test Set

featured an extended drawdown.

To see how the strategy holds up against the benchmark portfolios, the drawdown series are plotted and statistics are calculated. It is clear from the drawdown series, shown in Figure 17, that the strategy did not do as well avoiding drawdowns as it did in the validation series. However, it does feature lower drawdowns than the benchmarks quite often. The strategy is likely not as stable during the increased volatility in the test series compared to the validation set. Bias was also inevitably introduced in tweaking parameters to perform on the validation series, causing worsened performance on the test set by comparison.

The statistics for the strategy confirm decreased performance. The strategy did outperform the long SPY/short VXX portfolio and the short VXX portfolio in every category. It did not outperform the long SPY portfolio except in max drawdown. One of the objectives of the strategy is to minimize drawdowns, but the minimization of max drawdown is not in proportion to decreased returns compared to the long SPY portfolio.

Part of the reason the trading strategy had a poor overall performance was the big drawdown in February 2018. During this time, there was a sharp decline in major stock indexes, accompanied by a spike in the VIX. All the VIX ETFs needed to rebalance their positions by buying futures. The short ETFs had to buy to cover their short positions as they lost value, and the long ETFs had to buy more futures to rebalance after big gains. All the buying activity exacerbated the massive

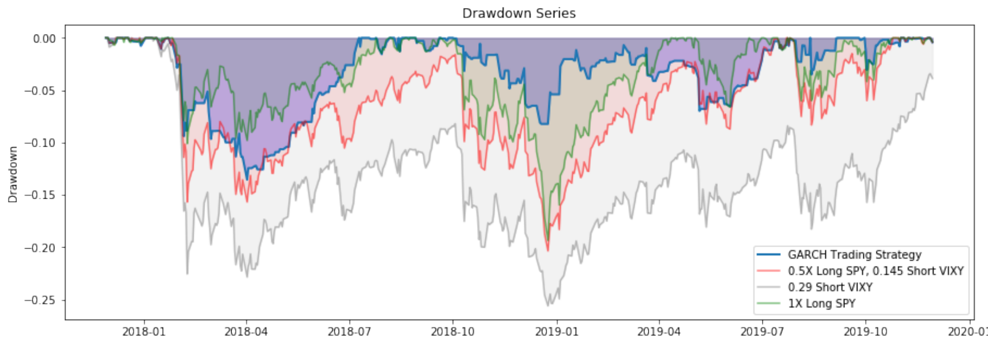

Figure 17: Trade Strategy and Baseline Drawdown Series on Test Set

Stat	GARCH Trade Strategy	Long SPY/Short VXX	Short VXX	Long SPY
Total Return	13.03%	12.92%	2.02%	24.25%
Daily Mean (ann.)	6.81%	7.52%	3.06%	11.97%
Daily Vol (ann.)	11.61%	16.88%	20.14%	14.92%
Daily Sharpe	0.59	0.45	0.15	0.80
Avg. Drawdown	-1.95%	-2.20%	-3.47%	-1.93%
Max Drawdown	-13.54%	-20.31%	-25.59%	-19.35%

Table 11: Trade Test Results

spike in the VIX. Since the trading strategy involves being short a VIX ETF, this obviously led to a steep drop, larger than the drop in the SPY.

This was a very low probability event, and in the time following the big VIX spike, the strategy outperformed all the benchmarks. The equity curve from April of 2018 (after the VIX futures markets had settled down) is displayed in Figure 18. It is clear the strategy does a good job avoiding drawdowns while matching returns. The drawdown series in Figure 19 visualizes the trading strategy's ability to avoid drawdowns.

The outperformance of benchmarks can be seen by the statistics in Table 12. The strategy matches the long SPY position in terms of total returns, has lower annualized volatility, a higher daily Sharpe, a lower average drawdown, and the max drawdown is less than half the SPY.

Though this is a cherry-picked time period, it shows that the strategy does have the ability to outperform the benchmark strategies.

6.3 Trading Strategy Conclusion

The trading strategy outperformed all benchmark portfolios in the validation set, but failed to do so on the test set. However, ignoring the improbable shock to the VIX futures market in February of 2018, the strategy would have outperformed the benchmarks. It appears that performance is better when overall market volatility is low. The validation set in 2017 featured very low general volatility and the strategy performed very well, but it did not perform as well under the high volatility in the test set. For comparison, the S&P 500 mean daily volatility was 0.31% in the

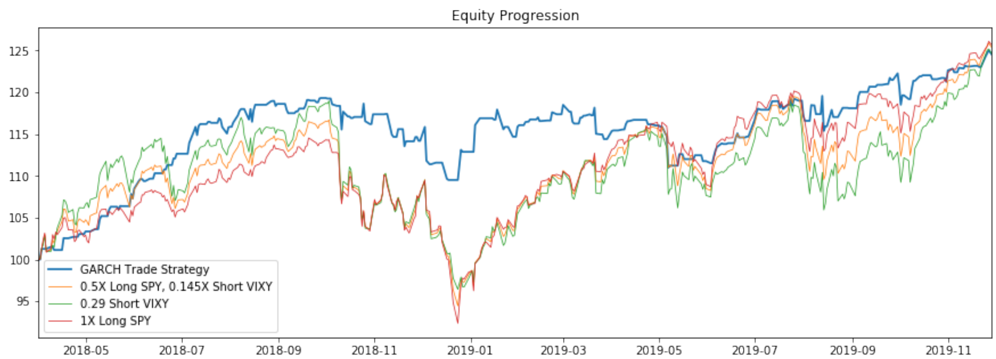

Figure 18: Trade Strategy Equity Curves on Test Set from April 2018 Forward

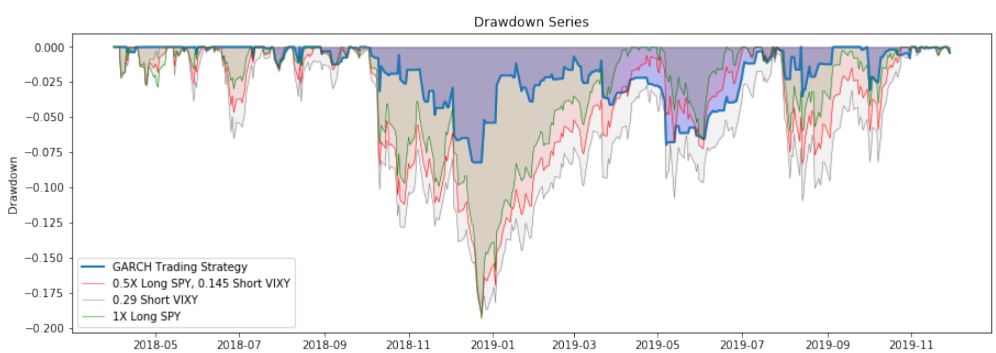

Figure 19: Trade Strategy and Baseline Drawdown Series on Test Set from April 2018 Forward

Stat	GARCH Trade Strategy	Long SPY/Short VXX	Short VXX	Long SPY
Total Return	24.50%	25.34%	24.58%	25.63%
Daily Mean (ann.)	13.63%	14.71%	14.73%	14.68%
Daily Vol (ann.)	10.08%	15.36%	17.70%	14.29%
Daily Sharpe	1.35	0.96	0.83	1.03
Avg. Drawdown	-1.13%	-2.04%	-2.48%	-1.78%
Max Drawdown	-8.21%	-19.05%	-18.94%	-19.34%

Table 12: Trade Test Results from April 2018 Forward

validation set at 0.65% in the test set, and respective medians of 0.19% and 0.45%. Likewise, the VIX experienced mean volatilites of 4.46% versus 6.11% and medians of 2.99% versus 4.47%. Visually, one can qualitatively see that when volatility is low in Figure 14 the equity curve in Figure 16 has heightened performance.

Additionally, refitting the model every so often could potential improve performance. The refit window sizes and refitting periods could be optimized and may improve the accuracy of the volatility predictions and trade signals. A simpler strategy that might achieve similar results would be to initiate the trades when volatility is below a rolling mean, and exit if the volatility rises above the rolling mean.

7 Conclusion

The GARCH model proved to be quite effective at forecasting volatility for the S&P 500 and the VIX. Simple (1,1) models were able to capture the trends of the volatilites. Moreover, the models were quite accurately able to predict increases and decreases in volatilities with accuracies in the 70%'s. This shows that GARCH is a very powerful tool and that while returns are not predictable, volatility can be predicted.

The trading strategy that was developed with the GARCH model proved that there is potential for the relationship between volatility and returns to be exploited. The trading strategy at times outperformed the market in metrics such as Sharpe ratio, average drawdown, and max drawdown. The strategy performs very well in periods of low volatility such as a sustained bull market, but performs poorly in choppy markets where volatility is more erratic.

There are additional applications of GARCH that were not explored in this report. These include Value at Risk which is a very powerful application of GARCH. Utilizing VaR forecasts a trader would be able to adjust his position sizes to market environments. There was also evidence of autocorrelation of returns in the PACF plots that was not explored.

Overall, GARCH is a powerful and versatile tool in a financial toolbox. In order to confirm the viability of the trading strategy, walk-forward testing needs to be done on true new data. Various tweaks to the model such as refitting may improve the strategy further and allow for the capture of upside in VIX and S&P 500 ETFs while minimizing return drawdowns. However, conclusions for the trading strategy are limited because of its mixed performance and does not show an ability to perform under conditions of high market volatility, and there are simpler models like rolling-mean-based methods that may very well match or exceed results of the strategy put forth in this report.