2020

Estimating Value-at-Risk of an Unconventional Portfolio

Elizabeth N. Mejía-Ricart
University of Richmond

Estimating Value-at-Risk of an Unconventional Portfolio

by

Elizabeth N. Mejía-Ricart

Honors Thesis

Submitted to:

Mathematics and Economics Departments
University of Richmond
Richmond,VA

April 30, 2020

Advisors: Dr. Saif Mehkari and Dr. Paul Kvam

1. Introduction

Since the 2008 financial crisis, interest rates and bond yields have been low all through the recovery and expansion that followed, and they are still low. As a result, more investors have been attracted to US equities, a space of possibly higher returns. However, these returns come with a potential downside: risk of loss. One of the methods to assess this potential downside is value-at-risk (VaR), which gained momentum in the late 1990s. At the time, the market risk amendment to the 1988 Basle Capital Accord required commercial banks with significant trading activities to put aside capital to cover market risk exposure to their trading accounts. VaR was used to determine the amount to be set aside (Lopez, 1998).

Formally, VaR is the maximum expected loss in a portfolio with a (1-θ)% confidence. This measure is developed to forecast the θ quantile of the profit-loss (P&L) distribution for a time period ahead (day, month, year). Much of the research done in this area has concerned the theoretical implementation of this method, and accuracy comparisons among calculation variations. At the same time, given the purpose of VaR's wide-spread use in the 1990s, the model has been mainly implemented to assess portfolios with short time-horizons and US-only exposure. For this reason, much of the little practical research in the topic has focused on comparing different VaR calculations in commercial banks against their daily P&L realizations. My aim is to test differently constructed VaR models using the holdings of the University of Richmond student-led ETF investment fund. The latter has exposure to non-US equities, making it a non-conventional case to test the accuracy of the models.

2. Literature Review

My research speaks to three papers in particular. Manganelli and Eagle (2001) propose a theoretical analysis of three different categories of estimating VaR. I use two techniques corresponding to some categories these authors propose (parametric and non-parametric) to calculate one-day ahead VaR and test which technique is empirically better. I use the RiskMetrics and historical simulation method and several GARCH techniques as my methods of comparison. I test the best empirical method by contrasting out-of-sample VaR estimations emanating from these models against the daily realizations of the ETF Fund's portfolio. Given the results brought by Manganelli and Eagle (2001), my thesis is that GARCH methods will outperform the RiskMetrics and historical simulation method.

The second paper that relates to this research performs a practical evaluation of different VaR methodologies. Berkowitz and O'Brien (2001) evaluate structurally-constructed one-day ahead VaR estimations for the trading revenues in six large commercial banks by juxtaposing this metric with the corresponding daily P&L realizations. Despite the detailed information employed in the bank models, their VaR forecasts did not out-perform forecasts based simply on an ARMA plus GARCH model of the banks' P&L. I will perform an analogous empirical study to analyze different VaR models of the ETF fund's holdings against the P&L daily realizations. I mimic the author's development of one-day ahead VaR metrics for the period of June 2018 to June 2019, using observations from the previous two years on a rolling basis. Moreover, instead of focusing on the performance of the same VaR metric for different funds — analogous to what the authors did — I compare the performance of different VaR metrics for the same fund. I also perform a similar analysis to the individual holdings of the ETF Fund, as I am interested in the models' response to greater volatility.

Lastly, another relevant paper for this research concerns the criticism of a particular specification of the VaR model. Simmons (2000) describes the shortcomings of VaR assuming a normal distribution, focusing on the fat-tails properties of financial markets. I expand upon this work by moving from the theoretical to the empirical. Looking at the realized volatility of the ETF fund's portfolio, I test the predictability of VaR by comparing the values obtained using different methods for calculating VaR against the realizations of the portfolio, using available information at time t. By using models that are less conservative in normality assumptions of the portfolio's P&L distribution, I empirically address the shortcoming concerns expressed by Simmons. At the same time, Simmons introduces the concept of tracking VaR in his research, or the X standard deviations of the portfolio's excess return compared to its benchmark, which is an interesting extension of VaR.

3. Data and Methods

The data used for this research was extracted from Yahoo Finance from September 2016 through June 2019 on the 20 current holdings of the University of Richmond ETF Investment Fund. These dates constraints were naturally imposed, given that September 15, 2016 was the inception date of FINX and BOTZ, two of the fund's holdings, and a third holding, INXX, was liquidated as of June 2019. Additionally, data on ACWI, the fund's benchmark, was extracted to illustrate the performance of the ETF's portfolio compared to its benchmark. Table 1 summarizes the tickers and full names of the ETF Fund's holdings, including ACWI.

Table 1: Summary of UR ETF Fund's Holdings and Benchmark Tickers and Descriptions

Ticker	**Name**
ACWI	MSCI All Country World Index
BOTZ	Global X Robotics & Artificial Intelligence ETF
ECH	iShares MSCI Chile ETF
EMQQ	The Emerging Markets Internet & Ecommerce ETF
EWN	iShares MSCI Netherlands ETF
EWZ	iShares MSCI Brazil Capped ETF
FINX	Global X FinTech ETF
HEWG	iShares Currency Hedged MSCI Germany ETF
HEWJ	iShares Currency Hedged MSCI Japan ETF
ICLN	iShares Global Clean Energy ETF
IHI	iShares US Medical Devices ETF
INXX	INXX Columbia India Infrastructure Index Fund
ITA	iShares U.S. Aerospace & Defense ETF
KIE	SPDR S&P Insurance ETF
VPU	Vanguard Utilities ETF
XLP	Consumer Staples Select Sector SPDR ETF
UUP	Invesco DB US Dollar Index Bullish Fund

Moreover, Table 2 presents data on daily prices for each of these holdings from September 2016 through June 2019, as well as the corresponding number of shares bought for each. Additionally, Figure 1 shows the fluctuations of the ETF Fund Holdings weighted by the number of shares obtained in the first place, against the performance of ACWI, assuming an equivalent number of shares of ACWI were bought at the beginning of the series.

Table 2: Summary of UR ETF Fund's Holdings Prices (Sept 2016 – June 2019)

	ACWI	BOTZ	ECH	EMQQ	EWN	EWZ
Min	56.74	14.7	36.07	22.74	23.17	30.72
1st Quartile	64.89	17.7	41.85	28.15	27.18	35.66
Median	69.83	20.02	44.44	32.36	30.19	38.7
Mean	68.31	20.18	45.14	32.56	29.26	38.65
3rd Quartile	72.55	22.95	48.32	37.1	31.46	41.82
Max	77.54	27.38	56.17	43.5	34.04	47.33
Shares	N/A	155	55	67	80	31
	FINX	**HEWG**	**HEWJ**	**ICLN**	**IHI**	**INXX**
Min	14.53	22.95	24.07	7.77	129.4	10.3
1st Quartile	17.87	25.94	28.92	8.59	159.3	11.76
Median	22.54	27.28	30.57	8.98	182.4	12.75
Mean	22.03	27.08	30.43	9.017	182.8	12.94
3rd Quartile	26.09	28.3	32.26	9.47	208.7	14.17
Max	29.31	30.33	35.21	10.5	237.7	16.6
Shares	61	96	46	169	7	111
	ITA	**KIE**	**VPU**	**XLP**	**UUP**	
Min	126.1	24.34	101.2	48.73	23.13	
1st Quartile	155.5	28.86	111.2	52.54	24.37	
Median	185.8	30.4	116.9	54.35	25.21	
Mean	179.2	29.94	116.6	54.07	25.11	
3rd Quartile	201.2	31.11	121	55.34	25.87	
Max	217.6	34.02	135.1	59.09	26.7	
Shares	8	46	9	36	51	

Figure 1: ACWI vs UR ETF Portfolio assuming equivalent initial investment

As it is perceived in the previous tables and graph, the ETF Portfolio has a diverse set of holdings, comprised of ETFs that have country-specific exposure, others that are more concentrated in a particular sector, and even a dollar hedge. Given that VaR was a measure that was designed for trading portfolios with US-only exposure, I am interested in evaluating its performance to this very different portfolio.

As noted previously, I pick two different techniques for calculating VaR and test which technique is empirically better. I estimate one-day ahead 5%[1] VaR for the period of June 2018 to June 2019 (estimation period) using the daily returns of the ETF Fund's daily portfolio from the previous 429 days. This implies that I use the returns from September 16, 2016 through May 31st 2018 to predict the one-day ahead VaR in June 1st, 2018. The window shifts one day to the right, from September 17, 2016 through June 1st, 2018, to estimate the next day's VaR, June 2nd, 2018. The method is generalized for all 263 VaR estimations.

Although the estimation window changes for every one-day ahead VaR prediction, a great portion of the information used to produce the VaR in the estimation period comes from the pre-estimation period, i.e. September 16, 2016 through May 31st, 2018. This merits further investigation regarding the similarity of the observations emanating from both periods. Table 3 summarizes the P&L returns distribution characteristics in the pre-estimation period, the estimation period and the whole 3 years (September 2016 through June 2019).

It is noteworthy that the three theoretical distributions possess average returns close to 0, however, both the pre-estimation and the added period possess means that are slightly negative. In terms of the standard deviation, the estimation period possesses the largest dispersion of data, and each standard deviation is significantly larger than the respective means. Moreover, the three timeframes possess skewness close to 0, indicating symmetrical returns, the pre-estimation period having the highest positive skewness. The main difference between the three theoretical distribution comes from the kurtosis, or tail characteristics. The estimation period and the overall timeframe possess tails that are less than 3 or slightly above 3, indicating less observations in the tails than the normal distribution, and similar tail behavior as

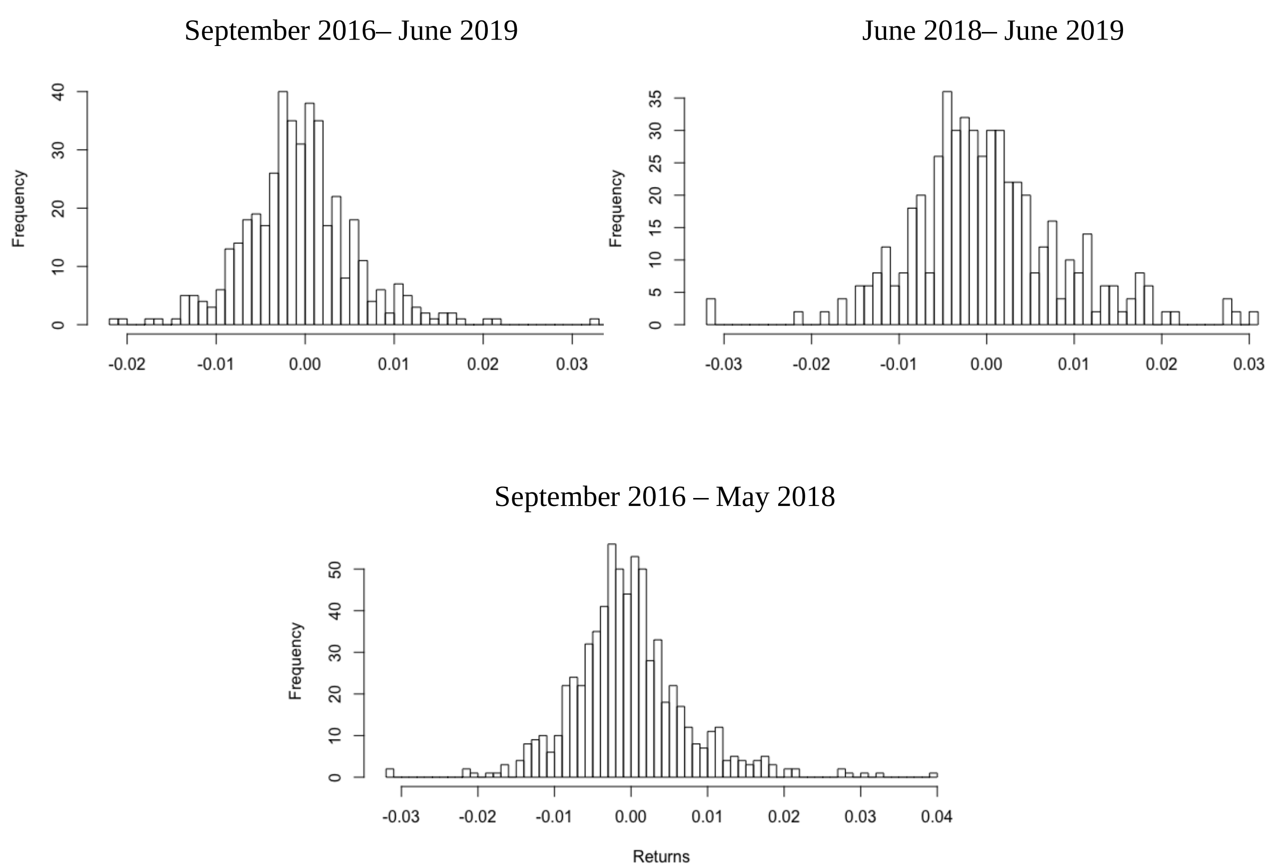
[1] This refers to the 5% quantile of the estimation.

the normal distribution, respectively. In contrast, the pre-estimation period has significantly fatter tails than the normal distribution. Figure 2 illustrates these empirical distributions graphically.

Table 3: Summary of UR ETF Daily P&L

Time Frame	Obs	Mean	Standard Deviation	Kurtosis	Skew
Sept 2016 - May 2018	429	-0.05%	0.66%	4.90	0.96
June 2018 - June 2019	263	0.01%	0.89%	1.60	0.37
Sept 2016 - June 2019	692	-0.03%	0.75%	3.17	0.66

Figure 2: Histograms of Daily Returns of ETF Fund Portfolio for Different Time Frames



3.1 Risk Metrics and Historical Simulation Hybrid Model

The first technique implemented to calculate one-day ahead VaR for the estimation period is the hybrid RiskMetrics and historical simulation (hybrid) method. This technique is non-parametric, that is, it makes no distributional assumptions about the estimation period. Instead, the maximum expected loss in this period is estimated by using the most recent K daily returns of the portfolio y_t, y_{t-1}, y_{t-K+1}, and assigning a weight $\frac{1-\lambda}{1-\lambda^k}$, $\left(\frac{1-\lambda}{1-\lambda^k}\right)\lambda, \ldots, \left(\frac{1-\lambda}{1-\lambda^k}\right)\lambda^{K-1}$, respectively. According to Table 4, the K used in this case was 429, as this is the number of observations used to predict VaR in the estimation period. Moreover, Boudoukh, Richardson and Whitelaw(1997) set λ equal to a value between 0.97 and 0.99, as no statistical method is available to estimate this unknown parameter. The theoretical difference in this range of numbers is that the lower end weights the most recent returns more heavily, while the upper end weights all returns more evenly.

Table 4: UR ETF Holdings Daily 5% VaR Estimation Summary

Specification	Num Obs for Estimation	Num Total Estimations	Mean VaR	Std. Dev. VaR
Hybrid				
λ= 0.97			-0.87%	0.09%
λ= 0.98			-0.90%	0.07%
λ= 0.99	429	263	-0.96%	0.09%
GARCH(1,1)			-1.32%	0.44%
IGARCH			-1.48%	0.56%

To get the VaR at the 5% level, these returns were organized in ascending order and their respective weights were summed until 5% was reached, starting from the lowest return. The VaR of the portfolio is the return corresponding to the last weight used in the previous sum. This estimation was done on a rolling basis to produce 263 one day ahead VaR estimates for the period of June 2018 through June 2019. Table 4 summarizes the results of these estimations. The VaR values produced for the one-year period in consideration were estimated using different values of λ. Each specification hybrid method produces similar results when looking at average one day ahead VaR estimates for June 2018 to June 2019, as well as the standard deviation of these predictions. Figures 3, 4 and 5 illustrate the VaR predictions against the daily P&L realizations of the estimation period. In theory, the P&L realizations should fall below the one-day ahead VaR estimates around 5% of the time. This will be evaluated in future sections.

Figure 3: One-Day Ahead Hybrid 5% Value-at-Risk (ƛ= 0.97) Against Daily Realizations

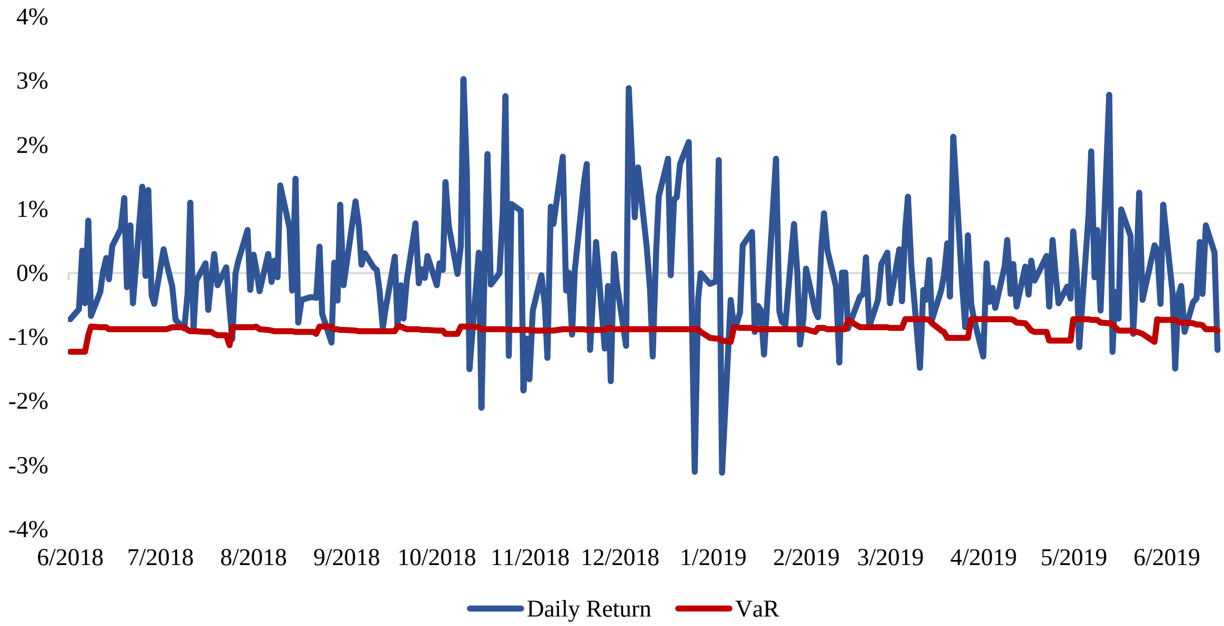


Figure 4: One-Day Ahead Hybrid 5% Value-at-Risk (ƛ= 0.98) Against Daily Realizations

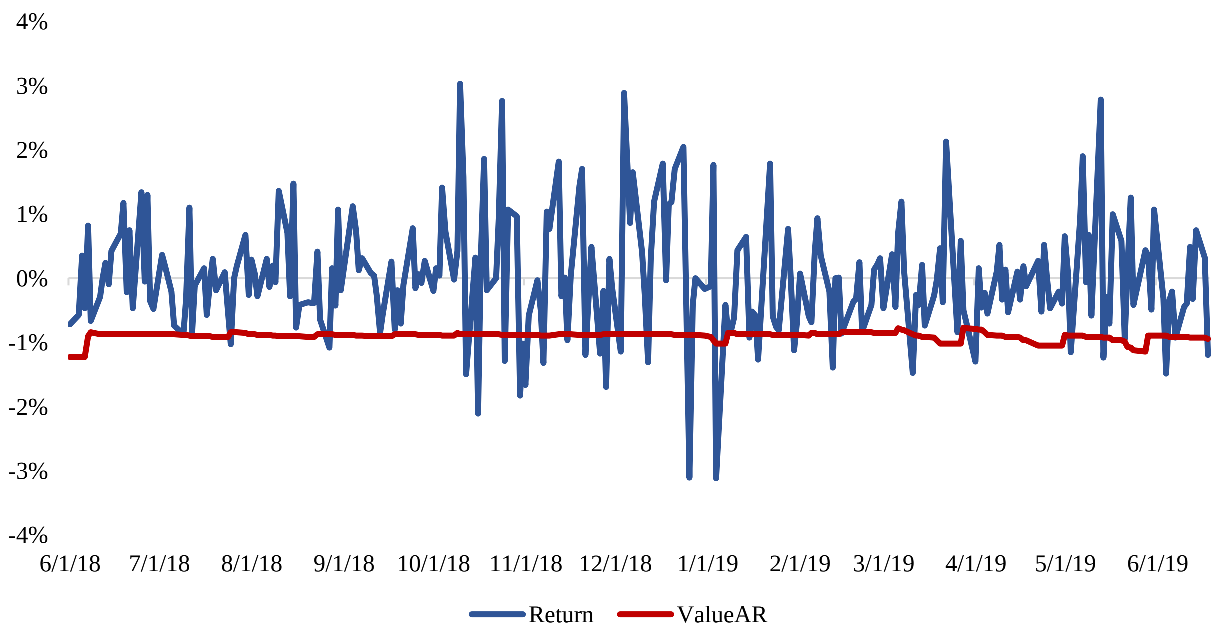
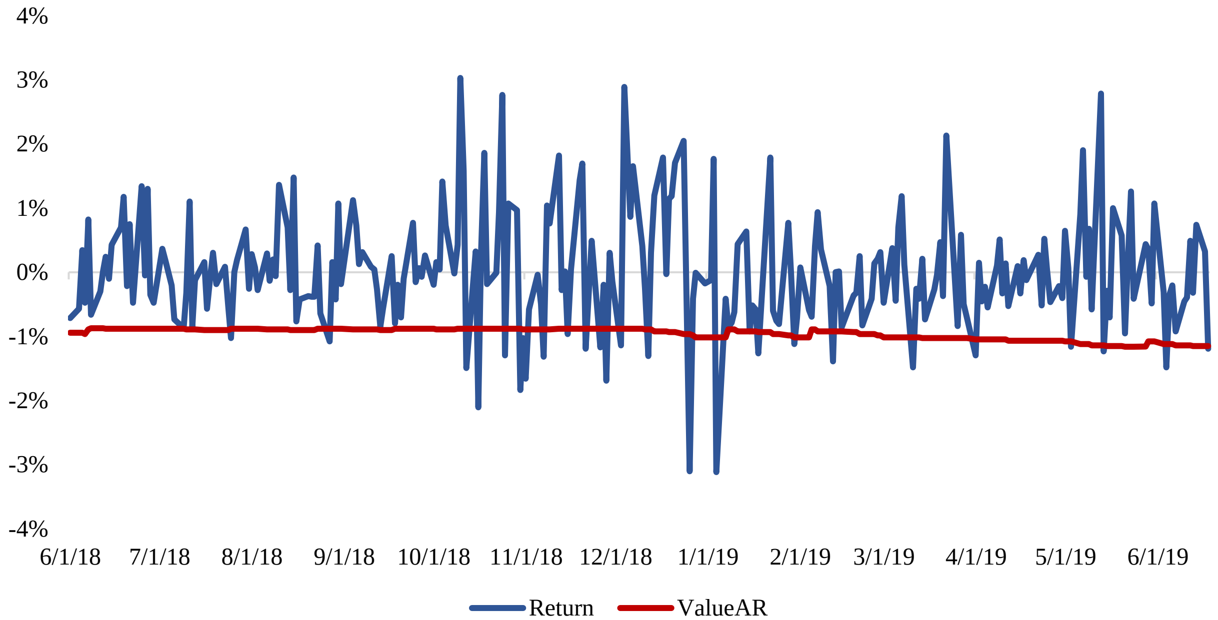


Figure 5: One-Day Ahead Hybrid 5% Value-at-Risk (ƛ= 0.99) Against Daily Realizations



3.2 Generalized Autoregressive Conditional Heteroskedasticity

The second approach used in estimating one-day ahead VaR for the ETF Fund portfolio is the GARCH(1,1) model. GARCH is a model for the realizations of a stochastic process imposing a specific structure of the conditional variance of the process. The GARCH(1,1) is the simplest specification of the model, which has become widely used in financial time series modeling and is implemented in most statistics and econometric software packages due to its relatively simple implementation (Williams, 2011).

Unlike the hybrid approach, the GARCH(p,q) model is a parametric model, that is, it makes assumptions about the profit-loss distribution to produce VaR estimates. In general, p refers to how far back the returns y_{t-p}^2 go and q refers to the order of the σ_{t-q}^2 portion of the estimator, which in this case are both 1. The GARCH(1,1) model makes the following distributional assumptions:

(1) $y_t = \sigma_t \varepsilon_t \qquad \varepsilon_t \sim N(0,1)$

(2) $\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2$

Where y_t is the returns at time t, σ_t is the standard deviation of the P&L distribution at time t, ε_t are randomly generated error terms drawn from the normal distribution and ω, α and β are unknown parameters that satisfy the second equation. A maximum likelihood estimation was performed to obtain the unknown parameters ω, α and β. Once the time series of estimated variance was computed, the standard deviation was multiplied by -1.645 to produce a 95% confidence VaR model. Just like the hybrid approach, 429 previous returns were used on a rolling basis to produce 263 one day ahead VaR estimates for the period of June 2018 through June 2019.

As seen in Table 4, the mean and standard deviation of all 263 VaR estimations are significantly larger in absolute value than that of any of the hybrid specifications. However, according to Figure 6, VaR estimates respond more drastically to the data than the hybrid method, which speaks to its accuracy. It is also noteworthy that the P&L realizations go below the VaR estimates less frequently than the hybrid model, as illustrated by Figure 6.

For sensitivity, a restricted version of the GARCH(1,1) model was also performed, the Integrated GARCH(1,1) or IGARCH(1,1) model, where the parameters α and β from equation (2), sum up to one. The following equation is used in place of (2):

(3) $\sigma_t^2 = (1 - \lambda)\, y_{t-1}^2 + \lambda\, \sigma_{t-1}^2$

Where λ is usually set equal to 0.94 or 0.97. As seen in table 3, and confirmed by Figure 7, the VaR values produced using the IGARCH(1,1) model are generally lower than both the GARCH(1,1) and Hybrid specifications. This is shown by the lower mean VaR exhibited by the IGARCH(1,1) model. However, there is also more volatility associated with this model, as indicated by the model's standard deviation.

Figure 6: One-Day Ahead GARCH(1,1) 5% Value-at-Risk Against Daily Realizations

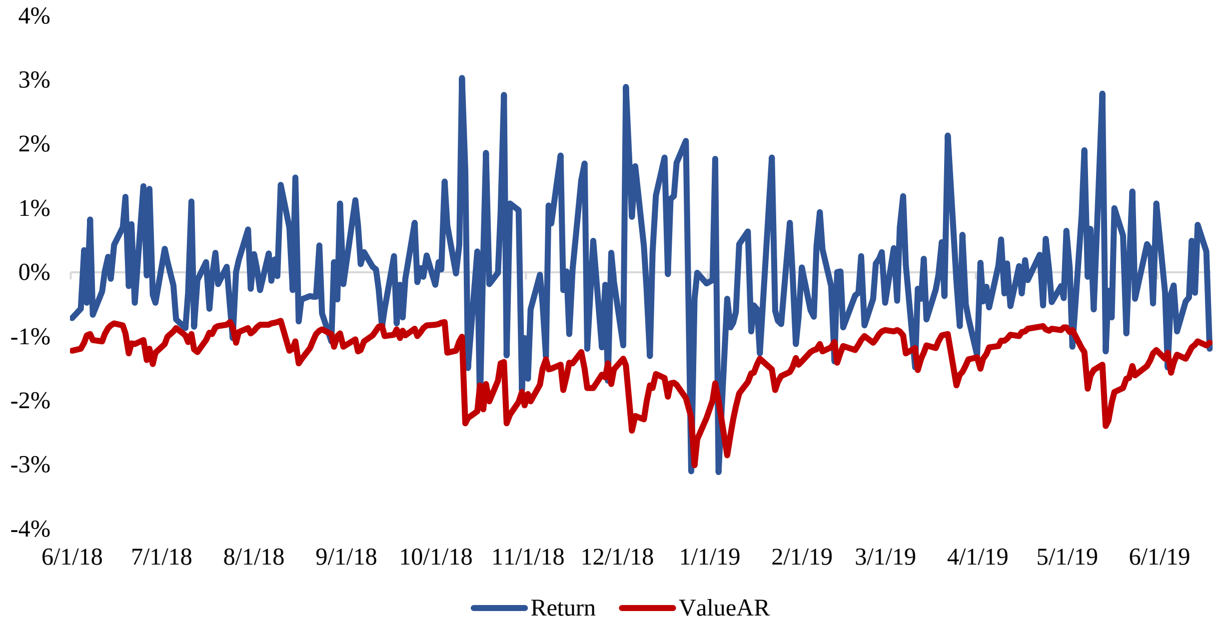
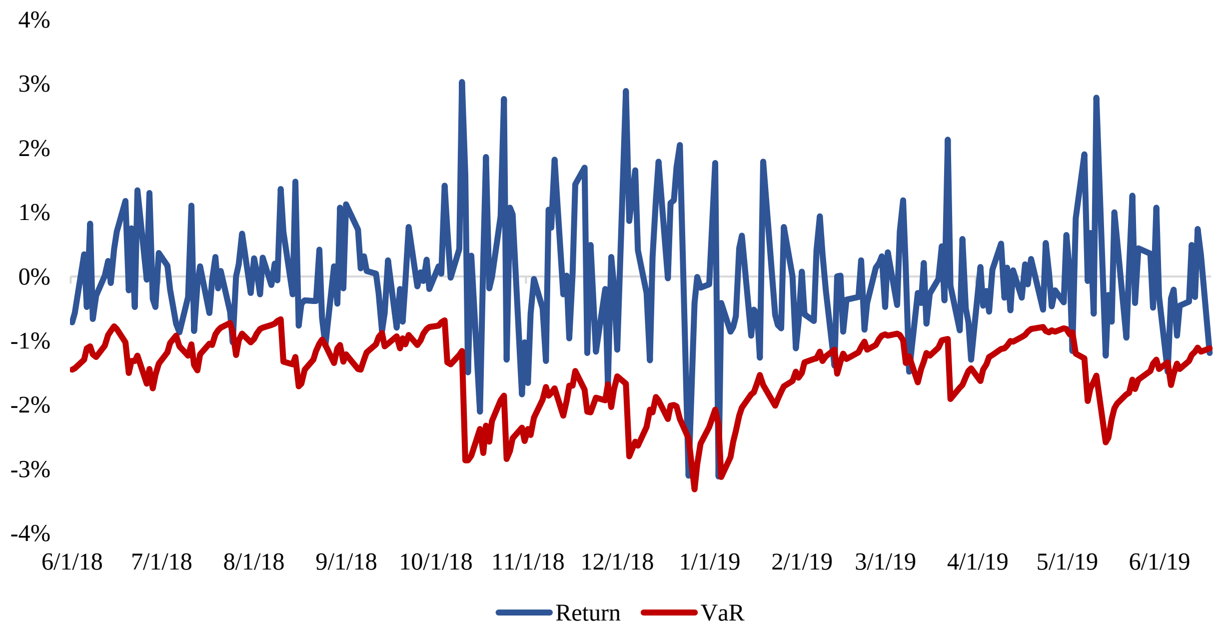


Figure 7: One-Day Ahead Integrated GARCH 5% Value-at-Risk Against Daily Realizations



3.3 Testing the Accuracy of the Models

The most relevant metric when evaluating the accuracy of VaR is the violation rate. A violation is defined by when a P&L realization are lower than the VaR estimation. As shown in Figures 3 through 6, this occurs when the blue line representing actual returns falls below the red line representing VaR predictions. If the VaR calculations are accurate, the percent of times that a violation occurs should be around 5%.

Table 5: UR ETF Holdings Daily 5% VaR Violations Summary

Specification	Total	Rate	Mean	Std. Dev	Min	Max	LR stat
Hybrid							
λ= 0.97	32	12.17%	0.51%	0.52%	0.01%	2.22%	20.7*, 21.9*
λ= 0.98	30	11.41%	0.51%	0.53%	0.00%	2.22%	16.9*, 17.7*
λ= 0.99	27	10.27%	0.50%	0.55%	0.00%	2.14%	11.9*, 12.2*
GARCH(1,1)	12	4.56%	0.34%	0.31%	0.02%	1.11%	0.74, 0.56
IGARCH	10	3.80%	0.26%	0.26%	0.01%	0.82%	0.86, 0.45

Table 5 summarizes the accuracy of the different models previously used to estimate VaR. As a first estimate, the number of total violations was calculated for each model. Furthermore, a violation rate was calculated based on the number of violations as a percentage of the total number of estimations. As mentioned before, this violation rate should be close to 5%. The violation rate in the three hybrid specifications is more than twice this percentage, with λ.= 0.99 having the lowest rate, at 10.27%. In contrast, the GARCH(1,1) VaR estimates are much more accurate. Moreover, the IGARCH model presents the lowest number of exceedances below the VaR values, only presenting 10 violations, 3.8% of the total estimated data points. In comparison, the number of violations and the violation rate with both GARCH models are less than half the size of the hybrid estimations.

It is also relevant to analyze the magnitude of these violations, that is, the difference between the VaR estimate and the actual return conditional on the return being less than the VaR estimate. For the hybrid approach, the specification with λ= 0.99 also seems to be the best performing estimate. The mean, standard deviation and the minimum violation are fairly similar across the three estimations. However, the model with λ.= 0.99 has a lower maximum violation. For the GARCH(1,1) model, the mean, standard deviation and maximum violations are smaller than the hybrid approaches, which is a result of the GARCH(1,1) having a smaller number of violations in general. Moreover, as a result of IGARCH's small number of violations, the mean, standard deviation, minimum and maximum violations are smaller than the hybrid and the GARCH(1,1) models.

Moreover, two hypothesis tests using likelihood ratio were conducted. The first one is Kupiec's proportion of failures coverage test (1995) which tests whether the observed frequency of VaR violations is consistent with the expected violations, given a chosen quantile (θ) and confidence level. Let x be the number of total violations in a given model, with n number of estimations, in our case 263. We treat x as a realization of a binomial random variable X, with probability of success of 5%. The hypothesis test in the unconditional coverage test is:

$$H_0: X \sim Bin(263, 0.05)$$

$$H_1: X \sim Bin(263, p)$$

The test statistic for this hypothesis test originates from a likelihood ratio, and the null hypothesis is rejected for large values of the statistic. In particular, H_0 is rejected when:

$$2log\left[\left(\frac{x}{263(0.05)}\right)^{x}\left(\frac{263-x}{263(0.95)}\right)^{263-x}\right] \geq \chi^2_{1,0.05} = 3.841$$

The left side of the inequality constitutes the test statistic for our experiment, while the right side constitutes the critical value for the test. The test statistics corresponding to the unconditional coverage test are reported as the first number in the last column of Table 5. An asterisk was placed next to the test statistics that are high enough to reject the null hypothesis.

A downside of the unconditional coverage test is that it does not consider any potential violation of the assumption of the independence of the number of exceedances. The conditional coverage test of Christoffersen et al. (1998) corrects this by jointly testing the frequency as well as the independence of exceedances, assuming that the VaR violation is modelled with a first order Markov chain. The test is also a likelihood ratio. Let i_t =1 refer to an instance when there is a violation at time t, and i_t =0 when there is not a violation at time t. The conditional coverage test has the following hypotheses:

$$H_0: \Pr(i_t = 0 | \, i_{t-1} = 1) = \Pr(i_t = 0 | \, i_{t-1} = 0)$$

$$H_1: \Pr(i_t = 0 | \, i_{t-1} = 1) \neq \Pr(i_t = 0 | \, i_{t-1} = 0)$$

The test statistic for this hypothesis test originates from a likelihood ratio, and the null hypothesis is rejected for large values of the statistic. In particular, H_0 is rejected when $-2log(\Lambda) \geq \chi^2_{1,0.05} = 3.841$, where Λ is the likelihood ratio test statistic. This number is reported as the second value in the last column of Table 5, and statistics yielding a rejection of the null hypothesis were also marked by an asterisk.

As seen in Table 5, all of the hybrid estimations yield high test statistics for the unconditional and conditional coverage tests, leading to a rejection of the null hypotheses. Hence, the hybrid models do not exhibit violation rates statistically similar to 5%, and these exceedances are not independent. In contrast, the small test statistics of the GARCH(1,1) and IGARCH VaR models for both hypothesis tests suggest a failure to reject the null hypotheses of these tests. That is, the observed frequency of these models' VaR violations is consistent with the expected violation rate and these violations are also independent from one another.

3.4 Modeling VaR of Individual ETFs

For the sake of sensitivity, the overall portfolio one-day ahead VaR estimates are compared against those of individual ETFs in the portfolio. Both the hybrid approach and the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) are used to identify whether these methods perform better for a portfolio compared to individual stocks.

3.4.1 RiskMetrics and Historical Simulation Hybrid

Table 6 shows the estimation results using the hybrid approach with ƛ= 0.99, a parameter that yielded the lowest number of violations for most ETFs compared to ƛ= 0.98 or ƛ= 0.97. It can be observed that the mean VaR tends to be lower for individual stocks than for the overall portfolio, attesting to the increased expected volatility of individual ETFs. The lowest average VaR was -2.76%, experienced by EWZ. Surprisingly, the volatility of the individual ETFs' VaR estimates, measured by their standard deviation was split evenly between being lower and higher than the overall portfolio's VaR volatility.

For the overall portfolio all of the hybrid models reject the null hypotheses of the unconditional and conditional coverage likelihood ratio tests. However, this is not the case for some ETFs. The null hypothesis for the unconditional LR test is that the frequency of VaR violations is consistent with the expected violations. In addition to this, the conditional LR test evaluates the independence of the number of exceedances. As shown by Table 5, for almost half of the ETFs VaR estimates, both null hypotheses fail to be rejected, as indicated by a low LR statistic. For further emphasis, the ETFs that have consistent expected violations and have independent number of exceedances are highlighted in gray. Interestingly, most of these ETFs – EWZ, ECH, HEWG, HEWJ and INXX – are single-country ETFs based outside of the United States. All of these international ETFs have exposure to a wide variety of economic sectors within these countries, contrary to the US-only ETFs held in the portfolio – ITA, XLP, VPU, FINX, KIE, IHI – which are mainly exposed to one sector in the US economy.

Furthermore, the international single-country ETFs have both lower mean VaR and higher standard deviation in VaR estimates compared to those of the overall portfolio. Although the violation rates for these ETFs were low, the violations size tended to be relatively large compared to their peers. The mean violation rate was as high as 1.23% for INXX, which also had one of the highest maximum deviations from VaR, 3.55%. The standard deviations of these violations were also larger than the average in most cases, which also speaks to the lower amount of exceedances in VaR estimations. These observations reveal that although the hybrid method with λ= 0.99 does a better job of producing VaR estimates for broad-exposure single-country ETFs, the magnitude of exceedances is less predictable. This speaks to the volatility of international markets, in particular emerging, and when their performance deviates from the mean, it does so dramatically.

Table 6: Daily 5% Hybrid (λ= 0.99) VaR Summary by ETF

ETF	VaR Estimation		Violations						
	Mean VaR	Std. Dev. VaR	Total	Rate	Mean	Std. Dev	Min	Max	LR stat
Portfolio	-0.96%	0.09%	27	10.27%	0.50%	0.55%	0.00%	2.14%	11.9*, 12.2*
BOTZ	-1.45%	0.08%	32	12.12%	0.85%	0.71%	0.01%	2.84%	20.7*, 21.0*
ECH	-1.69%	0.15%	13	4.92%	0.79%	0.76%	0.02%	2.58%	0, 0.2
EMQQ	-1.99%	0.18%	33	12.50%	0.90%	0.94%	0.01%	4.04%	22.6*, 22.6*
EWN	-1.11%	0.10%	29	10.98%	0.50%	0.45%	0.03%	1.80%	15.2*, 22.2*
EWZ	-2.76%	0.17%	17	6.44%	1.09%	1.10%	0.00%	3.48%	1.1, 1.1
FINX	-1.41%	0.10%	36	13.64%	0.77%	0.84%	0.00%	3.54%	29.0*, 30.1*
HEWG	-1.31%	0.05%	14	5.30%	0.77%	0.64%	0.14%	2.26%	0.1, 1.5
HEWJ	-1.37%	0.10%	19	7.20%	0.69%	0.72%	0.03%	2.93%	2.4, 5.4
ICLN	-1.43%	0.06%	20	7.58%	0.55%	0.50%	0.00%	1.47%	3.3, 3.4
IHI	-1.27%	0.06%	27	10.23%	0.70%	0.85%	0.00%	3.59%	11.9*, 12.7*
INXX	-1.72%	0.14%	20	7.58%	1.23%	0.99%	0.01%	3.55%	3.3, 3.4
ITA	-1.36%	0.07%	28	10.61%	0.55%	0.56%	0.01%	2.31%	13.5*, 15.2*
KIE	-1.11%	0.07%	22	8.33%	0.52%	0.60%	0.06%	2.59%	5.3*, 6.0*
UUP	-0.65%	0.09%	8	3.03%	0.19%	0.17%	0.00%	0.52%	2.5, 3.0
VPU	-1.07%	0.16%	26	9.85%	0.41%	0.27%	0.05%	1.09%	10.4*, 16.2*
XLP	-0.93%	0.09%	32	12.12%	0.39%	0.44%	0.00%	1.85%	20.7*, 20.7*

3.4.2 Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

Using the GARCH(1,1) and the IGARCH(1,1) models to estimate VaR for all individual ETFs in the portfolio yields estimates with correct number of exceedances and independent violations. That is, the one-day ahead VaR violations are not statistically different from 5% and they are independent from one another. Comparing Tables 5 and 6, it can be observed that the VaR estimates resulting from the hybrid model are relatively smaller in absolute value when compared to those estimated using the GARCH methods. This tradeoff between magnitude and accuracy is more than compensated by the latter models, given that the accuracy of the GARCH models are better, as measured by failure to reject both null hypotheses of the likelihood ratio test for both the overall portfolio and individual ETFs.

Considering Table 7 and 8, the overall portfolio has lower average VaR than most of the ETFs in the portfolio, when considered individually. This is to be expected, given that individual ETFs tend to be more volatile compared to the overall portfolio. The only exceptions to this behavior in our sample are XLP and UUP, the former being a defensive ETF focused in consumer staples and the latter being a dollar hedge, which is relatively stable over time. The most volatile ETF is EWZ, as measured by the most negative mean VaR using all the estimation method.

The largest number of exceedances corresponds to XLP, in both the GARCH(1,1) and IGARCH(1,1) models, with 19 violations and 16, respectively, which still yields a failure to reject both null hypotheses of the likelihood ratio. As expected, the mean deviation from the VaR estimates (the mean violation) is higher for most ETFs than the overall portfolio. Hence, generally the times that the profit-loss realizations go below the VaR estimations it does so more dramatically for individual ETFs. Moreover, the ETFs that have high mean violations, also have relatively high standard deviations, and maximum deviation from the estimated VaR. EWZ has the highest maximum violation in both models, as measured by the absolute value of the deviation between the realization and the GARCH VaR estimates.

Table 7: Daily 5% GARCH(1,1) VaR Summary by ETF

	VaR Estimation		Violations						
ETF	Mean VaR	Std. Dev. VaR	Total	Rate	Mean	Std. Dev	Min	Max	LR stat
Portfolio	-1.32%	0.44%	12	4.56%	0.34%	0.31%	0.02%	1.11%	0.74, 0.56
BOTZ	-2.27%	0.92%	13	4.92%	0.49%	0.44%	0.02%	1.65%	0.002, 1.36
ECH	-2.01%	0.35%	9	3.41%	0.75%	0.68%	0.10%	1.97%	1.54, 2.18
EMQQ	-2.75%	0.67%	15	5.68%	0.76%	0.75%	0.01%	2.72%	0.26,1.94
EWN	-1.50%	0.31%	12	4.55%	0.49%	0.34%	0.02%	1.11%	0.1, 1.26
EWZ	-3.37%	0.83%	10	3.79%	1.29%	1.42%	0.04%	3.70%	0.86, 1.66
FINX	-2.27%	0.96%	9	3.41%	0.55%	0.54%	0.07%	1.62%	1.54, 2.18
HEWG	-1.56%	0.33%	13	4.92%	0.49%	0.46%	0.10%	1.42%	0.002, 1.25
HEWJ	-1.75%	0.61%	8	3.03%	0.63%	0.55%	0.24%	1.55%	2.45, 2.96
ICLN	-1.63%	0.21%	15	5.68%	0.48%	0.45%	0.01%	1.38%	0.26, 1.96
IHI	-1.83%	0.70%	13	4.92%	0.49%	0.49%	0.02%	1.78%	0.002, 1.36
INXX	-2.19%	0.47%	13	4.92%	1.06%	0.89%	0.05%	2.92%	0.002, 1.35
ITA	-1.86%	0.58%	12	4.55%	0.39%	0.27%	0.02%	1.08%	0.11, 1.16
KIE	-1.44%	0.41%	15	5.68%	0.35%	0.37%	0.00%	1.38%	0.26, 2.09
UUP	-0.60%	0.04%	12	4.55%	0.16%	0.19%	0.02%	0.68%	0.11, 1.26
VPU	-1.43%	0.28%	9	3.41%	0.30%	0.20%	0.01%	0.58%	1.54, 2.18
XLP	-1.26%	0.38%	19	7.20%	0.26%	0.20%	0.02%	0.73%	2.42, 2.55

Table 8: Daily 5% IGARCH VaR Summary by ETF

	VaR Estimation		Violations						
ETF	Mean VaR	Std. Dev. VaR	Total	Rate	Mean	Std. Dev	Min	Max	LR stat
Portfolio	-1.48%	0.56%	10	3.80%	0.26%	0.26%	0.01%	0.82%	0.86, 0.45
BOTZ	-2.40%	1.00%	11	4.17%	0.44%	0.45%	0.02%	1.61%	0.39, 1.36
ECH	-2.09%	0.39%	9	3.41%	0.66%	0.67%	0.01%	1.82%	1.54, 2.18
EMQQ	-2.99%	0.74%	12	4.55%	0.72%	0.68%	0.04%	2.03%	0.11, 0.63
EWN	-1.67%	0.30%	8	3.03%	0.48%	0.20%	0.10%	0.76%	2.45, 2.96
EWZ	-3.72%	1.19%	8	3.03%	1.65%	1.53%	0.03%	3.81%	2.45, 2.96
FINX	-2.46%	1.05%	8	3.03%	0.43%	0.45%	0.01%	1.43%	2.45, 2.96
HEWG	-1.72%	0.43%	10	3.79%	0.47%	0.41%	0.04%	1.12%	0.86, 0.45
HEWJ	-1.87%	0.73%	8	3.03%	0.53%	0.54%	0.02%	1.38%	2.45, 2.96
ICLN	-1.71%	0.19%	13	4.92%	0.46%	0.40%	0.00%	1.11%	0.002, 1.36
IHI	-1.94%	0.76%	11	4.17%	0.45%	0.46%	0.03%	1.63%	0.39, 1.36
INXX	-2.38%	0.47%	13	4.92%	0.95%	0.72%	0.05%	2.64%	0.002, 1.36
ITA	-1.96%	0.61%	10	3.79%	0.38%	0.25%	0.13%	1.01%	0.86, 0.45
KIE	-1.58%	0.51%	11	4.17%	0.35%	0.33%	0.05%	1.19%	0.39, 1.36
UUP	-0.61%	0.05%	11	4.17%	0.15%	0.19%	0.04%	0.67%	0.39, 1.36
VPU	-1.48%	0.29%	8	3.03%	0.30%	0.23%	0.02%	0.54%	2.45, 2.96
XLP	-1.33%	0.43%	16	6.06%	0.28%	0.18%	0.00%	0.62%	0.61, 2.69

Conclusions and Future Research

A measure such as value-at-risk, which was designed for a trading and domestic portfolio can be effectively used to predict maximum expected loss for a portfolio that has international exposure. By using the portfolio of the University of Richmond's ETF Investment Fund, it was demonstrated that using the correct techniques for estimating VaR can yield accurate results. Non-parametric and parametric techniques were used to produce out of sample one-day ahead 5% VaR estimates and several analyses were performed to test the coverage and accuracy of these methods.

From all the techniques analyzed, the IGARCH(1,1) had the lowest violation rate. However, the VaR estimations produced by this method were not as close to the realizations compared the GARCH(1,1) model, suggesting that the accuracy of this model tends to be lower. The hybrid historical simulation and RiskMetrics approach did not yield significant results, as the violation rate or number of exceedances below the VaR estimate was significantly different from 5%.

The GARCH models produced the best performing estimates as measured by the violation rate, however, choosing between the GARCH(1,1) and the IGARCH(1,1) depends on the value placed on the accuracy of the model, as opposed to minimizing the violation rate. If the main goal is to have a fairly accurate estimate of VaR with a violation rate that fits into the desired confidence level, then a GARCH(1,1) should be used. On the other hand, if more weighting is placed upon having a lower violation rate, with higher than the desired confidence level, and with less regard for accuracy, the IGARCH(1,1) VaR model should be used.

Considering individual ETF holdings of the portfolio, the GARCH models once again dominate in performance overall. However, the hybrid approach improves its performance for broad-exposure single-country ETFs, although the magnitude of violations is less predictable. This improved performance for the latter ETFs is measured by failure to reject the unconditional and conditional coverage likelihood tests, indicating higher accuracy in the models. Additionally, when modeling individual ETFs VaR, we observe greater volatility reflected both in the increased size of average and standard deviation in the VaR, but also in greater mean and maximum violation sizes.

For future research, I recommend incorporating other optics to compare the accuracy of the VaR methods. In this paper, I have focused on the violation rate as being a crucial measure to evaluate these models. However, this analysis does not penalize models producing unnecessarily low VaR estimates. Moreover, Manganelli and Eagle (2001) suggested that semi-parametric models such as the conditional autoregressive value-at-risk (CAViaR) produce superior results compared to several parametric and non-parametric techniques. Incorporating this model into empirical research would be a great addition to the literature aimed at practical application and evaluation of VaR models.

References

Berkowitz, J., & O'Brien, J. (2002). *How accurate are value-at-risk models at commercial banks?*. The journal of finance, 57(3), 1093-1111.

Christoffersen, Peter (1998). Evaluating interval forecasts. *International Economic Review*, 39 (4), 841-862.

Ghalanos, A. (2019). *Introduction to the rugarch package*. Retrieved 13 December 2019, from https://faculty.washington.edu/ezivot/econ589/Introduction_to_the_rugarch_package.pdf

Kupiec, Paul H. (1995). Techniques for verifying the accuracy of risk measurement models, *Journal of Derivatives*, 3 (2), 73–84.

Lopez, J. A. (1999). *Methods for evaluating value-at-risk estimates*. Economic review, *2*, 3-17.

Manganelli, S., & Engle, R. F. (2001). *Value at risk models in finance*

Richardson, Matthew P. and Boudoukh, Jacob and Whitelaw, Robert F., The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk (November 1997). Available at SSRN: https://ssrn.com/abstract=51420 or http://dx.doi.org/10.2139/ssrn.51420

Simons, K. (2000). *The use of value at risk by institutional investors*. New England Economic Review, 21-21.

Williams, B. (2011). *GARCH(1,1) Models.*