

# Evaluating and Enhancing Value-at-Risk: A Regime-Switching Approach to U.S. Equity Risk

Junyi Guo

Fordham University, New York, NY, USA  
jguo120@fordham.edu

## Abstract

This paper evaluates the performance of Value-at-Risk (VaR) models under different market volatility regimes and proposes a regime-switching (RS) approach to improve risk measurement in U.S. equity markets. Using daily data from Apple, Microsoft, and the S&P 500 index, the study compares three traditional VaR methods: Historical Simulation (HS), Variance–Covariance (VCV), and Monte Carlo (MC). Market regimes are identified through rolling realized volatility, where periods of unusually high volatility are classified as turbulent. The study applies rolling one-day-ahead VaR forecasts and evaluates model performance through violation rates, Kupiec coverage tests, and breach severity analysis. The findings indicate that traditional VaR models generally perform adequately during stable market conditions but tend to underestimate downside risks during turbulent periods, particularly during events such as the COVID-19 pandemic. The proposed regime-switching VaR model combines the efficiency of the VCV approach during stable periods with the adaptability of the HS approach during turbulent periods. Results demonstrate that the RS method provides more stable coverage, fewer VaR violations, and lower breach severity across different asset types. The study highlights the importance of adaptive and regime-aware risk management models in capturing changing market dynamics and improving financial risk assessment.

**Keywords:** Value-At-Risk (Var), Regime-Switching Model, Financial Risk Management, Market Volatility, Historical Simulation, Variance–Covariance Method, Monte Carlo Simulation, U.S. Equity Market, Risk Measurement, COVID-19 Pandemic

## 1. Introduction

Value-at-Risk (VaR) is one of the predominant quantitative metrics for risks in the financial market, particularly in investment and portfolio management. In theory, VaR summarizes the maximum expected loss of a give portfolio over a given horizon at a specified confidence level (McNeil et al., 2015). For instance, a 1-day 99% VaR of 5% means that, under normal market conditions, the losses would exceed 5% only on 1 out of every 100 days. Due to its intuitiveness and interpretability, VaR is widely used by financial institutions and analysts to communicate risks to relevant stakeholders (Jorion, 1997).

However, despite its popularity, VaR is also heavily criticized for many of its drawbacks. Most notably, the stability of VaR is questionable and its performance is often poor under turbulent market times. Existing studies such as Shaik & Padmakumari (2022) suggested that VaR has a tendency to underestimate tail-end risks during notable market events such as COVID-19 pandemic. Danielsson et al. (2018) discussed the impacts of financial crises on volatility and noted that measures of VaR can become unreliable during disruptions. The actual maximum losses often surpass prescribed losses by VaR due to a mixture of factors include violation of assumptions and lack of forward-looking mechanisms (Tolikas, 2014). These downsides have raised significant questions on the reliability of VaR.

One key limitation from existing studies in the field is that most studies assume volatility of the market is stable over time. In reality, market volatility is often cycling through high and low volatility periods due to various macroeconomic and geopolitical factors. This is referred to as market regime changes. VaR model tends to perform better in more stable markets and struggles with sudden change in market volatility (Ranković, 2016). These characteristics further put doubts on the reliability of VaR as a leading risk management measure in financial markets.

The objective of this paper is to investigate whether performance of VaR is regime-dependent and whether a simple regime-switching rule can improve risk measurement. Based on data collected on two large US stocks (Apple and Microsoft) and one broader market index (S&P 500), this paper calculates VaR by different methods and benchmarks them against actual returns. Market regimes are identified using rolling realized volatility, with periods of unusually high volatility classified as turbulent. This paper further proposes a regime-switching VaR method by changing methods based on external market environment. The resulting method yields an improved performance and can better capture sudden volatility spike and reflect the trend in VaR.

The rest of this paper is organized as follows. The second section reviews advantaged and limitations of VaR methods from existing literature. The third section outlines data collection, sampling strategy, and analysis methodology. The fourth section presents and discusses results. The last section concludes this paper.

## 2 Literature Review

Typically, VaR can be calculated through three approaches which are the historical simulation method, variance-covariance

method, and Monte-Carlo simulation method.

The historical simulation (HS) method estimates VaR directly from historical returns and use percentile distribution to sample VaR. This method thus bypasses some important assumptions such as normality of returns which is also its main advantages (Pritsker, 2006). However, the major limitation is that this method also depends on historical return window selection properly. It is generally slow to adapt to changes in market volatility which can happen often in real markets.

The variance-covariance (VCV) method, on the other hand, takes a parametric approach and uses key distribution metrics (mean and standard deviation) to compute VaR. Unlike HS method, VCV method centers around the normality of returns of the securities. It is easy to compute due to its simplicity and is often used in large portfolios and regulatory reporting for quick and effective risk interpretation (Cont, 2001). However, empirical evidence generally suggests that financial returns are not normally distributed most time. They usually exhibit significant left-tail risks, especially during period of high turbulence (Abad et al., 2014). This creates systematic underestimation of tail risk for distribution-based measures (Engle, 1982).

Furthermore, the Monte-Carlo simulation (MC) approach simulated returns based on assumed distribution and then calculate VaR accordingly. In theory, the MC approach is the most flexible option and accounts for imperfections such as non-normal distribution and time-varying volatility. It can also incorporate additional new information and make return forecasts path dependent. Nonetheless, these additional features also increase the complexity of this method and actual performance depends heavily on the data-generating strategy (Glasserman et al., 2000). Simple MC approach tends to replicate results from VCV method.

Therefore, each method has its own advantages and limitations which will be examined in the literature review section. Generally, VCV is computationally efficient while HS is more lenient on assumptions. MC is more flexible but requires additional control on quality and accuracy of assumed distribution of returns. Misspecification in assumptions can significantly affect estimate accuracy and reliability (Dowd, 2005).

As a popular risk management metric, a large number of literatures has shed lights on back-testing VaR. For instance, Kupiec (1995) proposed the unconditional convergence test which compares observed violation of VaR to nominal confidence level. This method is adopted in this study to provide benchmarking results and contribute to the main research objectives. In addition, many studies also pointed out that VaR is not performing as expected during market turbulence. For instance, studies such as (Shaik & Padmakumari, 2022) have documented VaR tend to yield high violation rate during the COVID-19 pandemic and this is particularly evident for parametric-based approach. The general consensus is that simple VAR design is not suitable to handle the increasing market complexity but more sophisticated models also required fine-tuned inputs and lacks uniformed approaches. Furthermore, studies such as Artzner et al. (1999) also critiqued the lack of subadditivity in the theoretical space of VaR.

**Table 1** provides a summary of key results.

**Table 1:** Advantages and Limitations of Different VaR Methods

Method	Core Idea	Main Strength	Main Limitation	Regime Performance
Historical Simulation (HS)	Uses empirical distribution of past returns	Captures fat tails and extreme losses	Slow to adjust when volatility falls	Performs best in turbulent markets
Variance–Covariance (VCV)	Assumes normally distributed returns	Efficient and stable in calm periods	Fails under fat tails and volatility clustering	Performs best in stable markets
Monte Carlo (MC)	Simulates returns from assumed distribution	Flexible framework for complex assets	Inherits distributional errors (e.g. normality)	Similar to VCV
Regime-Switching (RS)	Uses VCV in stable regimes and HS in turbulent regimes	Adapts to changing market conditions	Requires regime identification	Robust across regimes

### 3 Data and Research Methodology

This paper collects stock data of US listed companies Apple and Microsoft, and the major market index S&P 500 from Yahoo Finance. The two stocks are large-cap tech stocks and the index is often used to approximate the broader market performance (Bodie et al., 2015). They are selected to assess whether results can be generalizable across different equity types.

The daily adjusted closed price is collected from the period of 2015 to early 2016. This period should cover various market regime changes, most notably, the COVID-19 pandemic and subsequent volatility episodes and inflation. This long period ensures there are enough market volatility cycles in the period to capture the VaR characteristics of different methods.

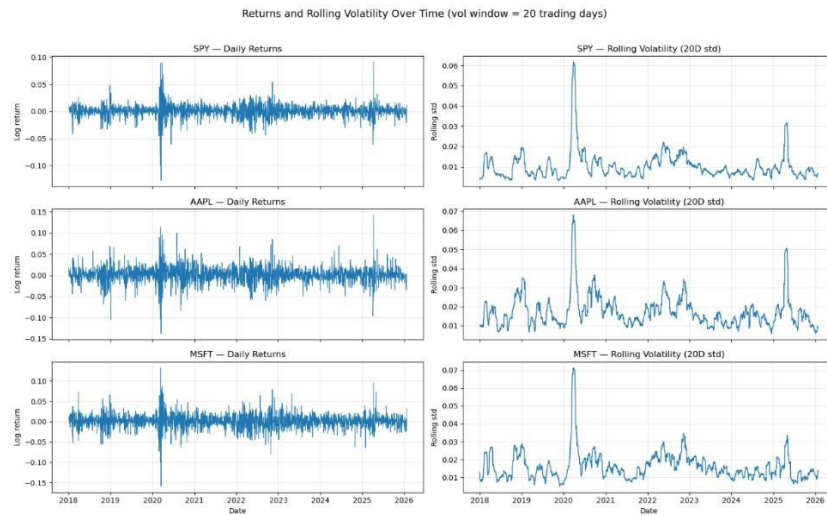
The daily returns are calculated with daily log returns specified below:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

Figure 1 presents some information on the returns and volatility of the three securities. All of them seem to exhibit the same general trend, with roughly the same spike and trough timing. For instance, all the stocks show sharp volatility jumps in early 2020 in response to the COVID-19 pandemic, and individual stocks tend to be more volatile than market index which his

expected.

**Figure 1: Returns and Volatility of the Three Securities**



These returns are then used to compute VaR by different methods and with rolling windows to generate out-of-sample one-day-ahead risk forecasts. VaR is always presented as a positive number to indicate maximum losses of a given confidence interval. Specifically, the three different VaR approaches are illustrated below. All models use 99% confidence interval for calculation purposes.

**Historical Simulation (HS)**

The HS approach uses percentile of historical daily returns calculated above over the last 250 trading days. The formula is shown below.

$$VaR_{t,Historical} = -Q_{\alpha}(r_{t-250}, \dots, r_{t-1})$$

This approach, as reviewed in section 2, does not depend on distribution of historical returns but is generally slow to adjust to new information.

**Variance–Covariance (VCV)**

The VCV approach is parametric-based and assumes normality of returns.

$$VaR_{t,VCV} = -(\mu_t + z_{\alpha} * \sigma_t)$$

The mean and standard deviation is again based on the distribution characteristics of the last 250 trading days.

**Monte Carlo (MC)**

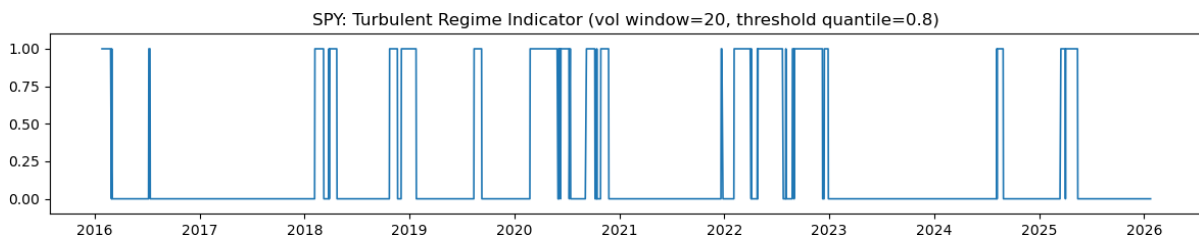
The MC approaches simulate 50,000 returns following a normal distribution with the same distribution metrics obtained from the VCV approach. In practice, performance of MC and VCV approach should be close, unless MC has additional customized paths.

Another important concept in this paper is market-regime switch. This is identified by setting a threshold based on past volatilities. Specifically, if rolling volatility exceeds 80% percentile of volatility series over the full sample, the days to compute the current rolling window are categorized as turbulent while other days as normal. The look back window is 20 days. This is a common approach and thresholds to identify volatility regimes established under studies such as Hamilton & Susmel (1994).

$$D_t = \begin{cases} 1, & \text{if } \sigma_t > k \\ 0, & \text{if } \sigma_t \leq k \end{cases}, \quad k = Q_{0.8}(\sigma_{t-20}, \dots, \sigma_{t-1})$$

Figure 2 shows the mapping of turbulent regimes by year. The results indicate that the turbulence is largely consistent with known market volatility drivers such as COVID-19 pandemic and increasing geopolitical tensions.

**Figure 2: Market Volatility Regimes During the Sample Period**



This paper also proposes a new regime-switching (RS) VaR. Specifically, the model uses:

1. VCT approach during stable time periods. This can take advantage of the simpleness and efficiency of the method

under the stable market environment.

2. HS approach during turbulent time periods. This can better address the non-normality and fatter tails of returns.

$$VaR_{t,RS} = \begin{cases} VaR_{t,VCV}, & \text{if } \sigma_t \leq \sigma_{80\%} \\ VaR_{t,HS}, & \text{if } \sigma_t > \sigma_{80\%} \end{cases}$$

To back-test performance of different VaR methods, this paper employs three approaches. The first one is violation rate, which calculates the proportion of days that VaR is less than the actual losses. The violation rate is calculated based on a 60-day rolling window. Out-of-sample is generally considered to be best practice when back-testing model performance to avoid overfit (Hansen, 2005). The formula for rate is below.

$$I_t = 1(r_t < -VaR_t)$$

The second method is the Kupiec unconditional coverage test. It takes on the violation rate and further tests whether they are consistent with the nominal confidence level. It follows the method under Kupiec (1995) and significance is tested through chi-squared tests.

The last method is breach severity. It calculates the magnitude of losses beyond the VaR suggests. It can be expressed as:

$$S_t = |r_t| - VaR_t$$

This is only calculated on days with violations. Together, these three methods should provide a holistic review on how often and severe VaR underestimate risks.

### 4 Results and Discussion

Figure 3 shows the cumulative violations of the three common VaR models and the RS customized approach designed by this paper. The results show that number jumps and divergence happen most often during market turbulence as suggested in figure 2. In S&P 500 and Microsoft, VCV and MC approaches yield nearly identical results while some noticeable difference exist in Apple. Out of all models, HS approach tends to perform best with the smallest number of total violations. It accumulates violations more slowly compared to the other two methods. The RS method, meanwhile, provides a performance between VCV and HS. It offers worse but still acceptable performance during stable times, and is quicker to adapt to new volatility pattern during turbulent times.

This overall leads to less total violations than VCV and MC while offering more stable performance over the entire time period. The benefits of RS approach is consistent across all three assets. This shows that this benefit is not asset or type dependent. It can be generalizable.

**Figure 3: Cumulative VaR Violations**

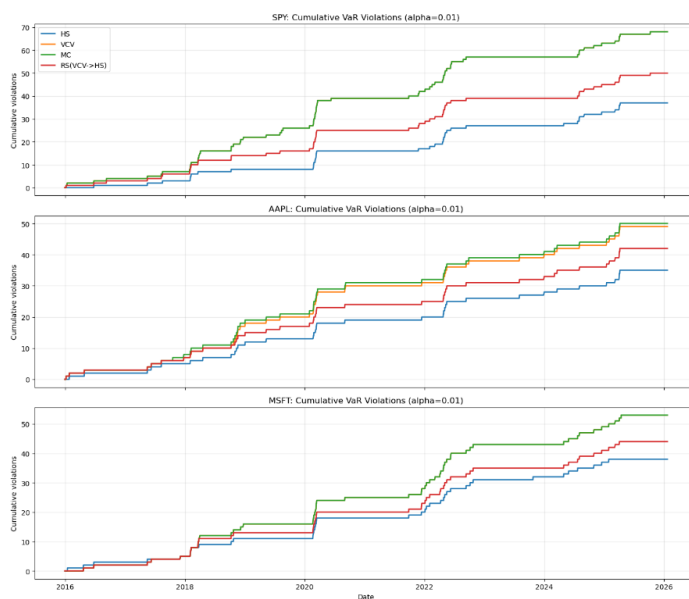


Figure 4 shows the rolling violation rates of Apple, Microsoft, and S&P 500 over the years by the RS method discussed in section 3. The results reiterate that violations tend to coincide with turbulent periods and market-regime changes depicted in figure 2. For instance, volatility rose sharply in early 2020 and rolling violation rates of the three stocks all exceed 10%, the highest of the sampling period. This reflects a market-level panic in response to the pandemic (Zhang et al., 2020) Out of the three securities, RS approach violation rates of most times during the sample period are well under the 1% confidence level. This indicates VaR approaches generally perform well during table times. However, during turbulent times, VaR violation rates rose quickly and would return to normal after a few rolling cycles. This plot shows that the RS method

can provide stabilization impacts on VaR violation and help it reconcile quickly after experiencing sudden spikes in volatility. Again, this pattern is consistent across Apple, Microsoft, and S&P 500.

**Figure 4: Rolling Violation Rate**

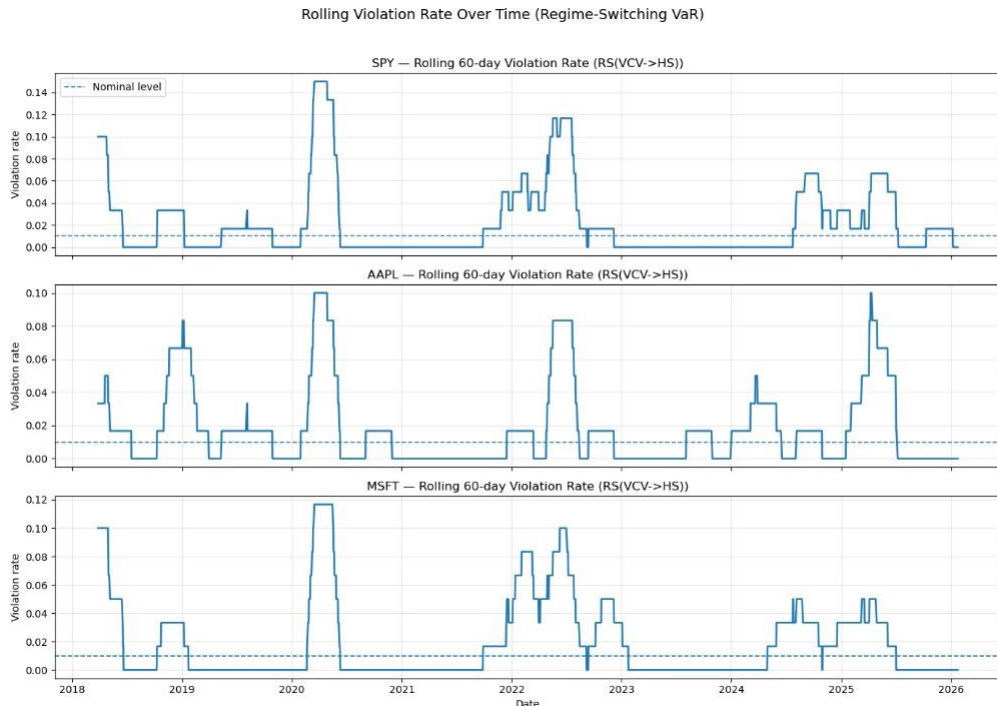


Figure 5 provides the Kupiec coverage test results. For S&P 500, HS method tends to pass the test in stable time periods, while all other methods tend to fail. The results suggest systematic underestimation of risks by VCV and MC methods under stable time. During turbulent times, all three VaR methods yield weak coverage. For Microsoft and Apple, the individual stocks, none of the VaR methods can provide significant coverage during both stable and turbulent times. In some instances, the p-value is closer to the confidence level but still above them slightly. Volatility spikes and fat-tailed distribution can thus significantly weakened VaR validity during market turbulence (Likitratcharoen & Suwannamalik, 2024).

The patterns confirmed two findings. First, VaR performance is strongly related to volatility regimes. For different stocks, the methods tend to work well for one regime (stable or turbulent) and perform poorly for the other one. The methods have their strengths and weaknesses. No single traditional VaR method can handle all market conditions. This provides some justification to the proposed RS method. Second, during turbulent market times, the returns are mostly not normally distributed and often exhibit extreme values of left tails (McNeil et al., 2015). The volatility clustering amplifies tail-end risks and affects reliability of traditional methods and render them largely unsuitable. During market stress, the percentile/cut-off-based approach of VaR can no longer properly depict the actual risk patterns due to the distorted distribution. This highlights an existing limitation of VaR and also calls out for more adaptive methods to capture the changing market conditions more quickly.

**Figure 5: Kupiec Coverage**

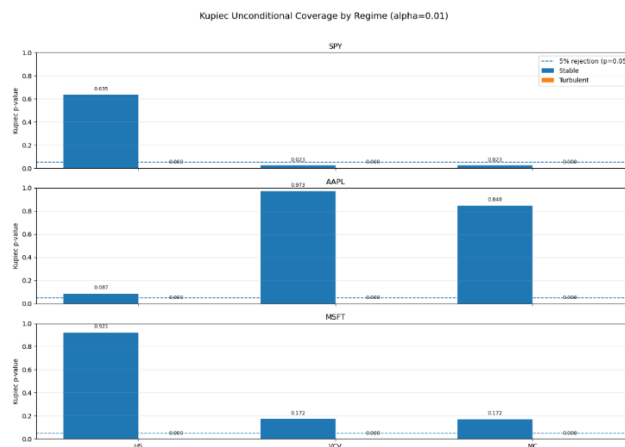


Figure 6 shows the breach severity of RS method on the three securities. The major breach occurs during the COVID-19 pandemic and breaches in other time periods are limited and usually security specific. Individual stocks, even though large

cap, tend to have more volatility spikes and thus more breach severity than the market index S&P 500. This shows the RS method can provide proper and adaptive coverage of risks than conventional VaR, whose performance are not ideal as shown in figure 3 and 5. In stable markets, the RS methods perform slightly worse than normal VCV methods but the performance deterioration is not concerning. The breach severity is generally very low for RS method during stable market times. In turbulent market times, the RS method can more quickly capture the sudden volatility spike and reflect hem more properly in VaR. This trade-off is advantageous since it provides more down-side protection while sacrificing performance slightly during non-risky periods. The losses mostly happen during the traditional phases (e.g., from turbulent to stable), but the overall scale is very limited.

Nonetheless, it is still important to point out that the RS method is still not capable at handling extreme volatility. This is more of an inherent limitation of VaR and requires more complex solutions such as integration with other more sophisticated forward-looking methods (Abad et al., 2014). This behaviors usually occurred during extreme market stress (e.g., top 5% of volatility) and the RS method shows acceptable breach severity under most other market turbulent periods measured under figure 2.

**Figure 6: Breach Severity**

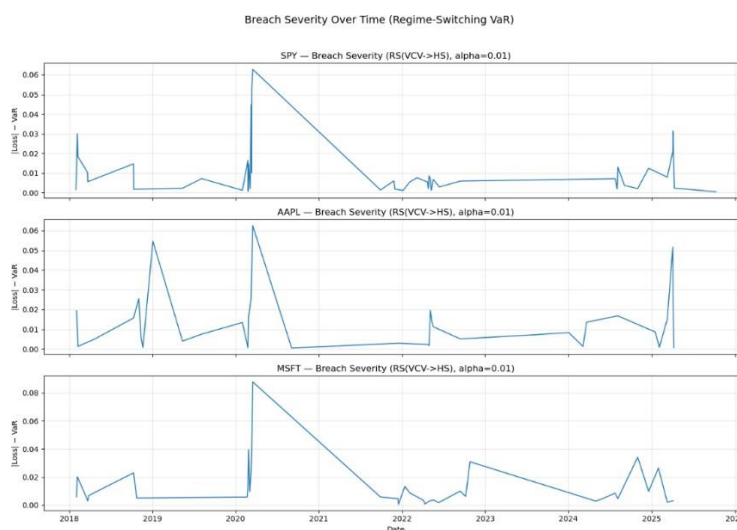
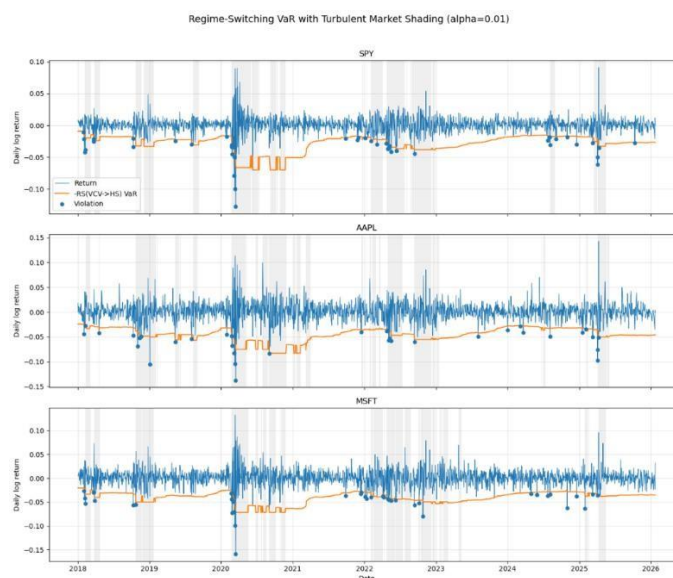


Figure 6 shows various metrics of RS method VaR and how they behaved during market turbulence. The shaded region identifies market turbulence and returns, actual VaR, and violation rate are all shown in the figure.

**Figure 7: Regime-Switching VaR Benchmarking Against Market Turbulence**



RS plot shows an intuitive comparison between RS method VaR and actual returns during turbulent times. In most cases, the RS VaR can properly capture downside risks of the three securities. The RS VaR can quickly capture changes in volatility patterns. For instance, in early 2020, as COVID-19 pandemic hit, the RS VaR threshold rose sharply as the model switches from VCV to HS which better captures the tail-end risks. Thus, most extreme losses are closer to RS VaR and the adaptive mechanisms provide the major contribution here.

Outside the market turbulence, the RS method remains smoother and less conservative. This allows faster and simple calculation during stable times where risk is low. The violations of different securities tend to cluster in the same market turbulence periods. This suggests the regime-dependent nature of tails risks and their implications on VaR. The findings thus support an adaptive VaR method based on prevailing volatility patterns.

## 5 Conclusion

To sum up, this paper collects US stock data to analyze performance of VaR and evaluates their connection to market volatility regimes. The findings suggest that VaR tend to perform reasonably during stable markets but fail to capture downside risks during market turbulence. To address such limitations, this paper proposes an adaptive VaR model based on market volatility regimes and the results yield improving performance. The RS VaR method combines benefits of HS and VCV methods and adds an adaptive mechanism. It can still identify risks effectively during table markets and quickly capture the sudden volatility jumps during market turbulence. The results in this paper show that RS VaR can yield more stable coverage, lower counts of violations, and lower breach severity as well. Though it still has trouble modeling risks during extreme market events, it can provide meaningful improvement to simple methods such as HS and VCV. These findings thus support the use of adaptive, regime-aware risk models for more effective financial risk management. Some examples would be GARCH-based risk models, which tend to be better positioned at capturing the complex dynamics of the market.

## References

- [1] Abad, P., Benito, S., & López, C. (2014). A comprehensive review of value at risk methodologies. *The Spanish Review of Financial Economics*, 12(1), 15–32.
- [2] Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228.
- [3] Bodie, Z., Kane, A., & Marcus, A. J. (2015). *Portfolio management*. McGraw-Hill Education.
- [4] Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4), 841–862.
- [5] Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236.
- [6] Danielsson, J., Valenzuela, M., & Zer, I. (2018). Learning from history: Volatility and financial crises. *The Review of Financial Studies*, 31(7), 2774–2805.
- [7] Dowd, K. (2007). *Measuring market risk* (2nd ed.). John Wiley & Sons.
- [8] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.
- [9] Glasserman, P., Heidelberger, P., & Shahabuddin, P. (2000). Efficient Monte Carlo methods for value-at-risk. *Mathematical Finance*, 10(4), 437–457.
- [10] Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64(1–2), 307–333.
- [11] Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics*, 23(4), 365–380.
- [12] Jorion, P. (1997). *Value at risk: The new benchmark for managing financial risk* (Vol. 2). McGraw-Hill.
- [13] Kupiec, P. H. (1995). *Techniques for verifying the accuracy of risk measurement models* (Working Paper No. 95–24).
- [14] Federal Reserve Board.
- [15] Likitrachoen, D., & Suwannamalik, L. (2024). Assessing financial stability in turbulent times: A study of generalized autoregressive conditional heteroskedasticity-type value-at-risk model performance in Thailand's transportation sector during COVID-19. *Risks*, 12(3), 51.
- [16] Longin, F. M. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking & Finance*, 24(7), 1097–1130.
- [17] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools* (Rev. ed.). Princeton University Press.
- [18] Pritsker, M. (2006). The hidden dangers of historical simulation. *Journal of Banking & Finance*, 30(2), 561–582.
- [19] Ranković, V., Drenovak, M., Urošević, B., & Jelić, R. (2016). Mean-univariate GARCH VaR portfolio optimization: Actual portfolio approach. *Computers & Operations Research*, 72, 83–92.
- [20] Shaik, M., & Padmakumari, L. (2022). Value-at-risk estimation and backtesting during COVID-19: Empirical analysis based on BRICS and U.S. stock markets. *Investment Management & Financial Innovations*, 19(1), 51–64.
- [21] Tolikas, K. (2014). Unexpected tails in risk measurement: Some international evidence. *Journal of Banking & Finance*, 40, 476–493.
- [22] Zhang, D., Hu, M., & Ji, Q. (2020). Financial markets under the global pandemic of COVID-19. *Finance Research Letters*, 36, 101528.