Banking on Uncertainty: An Analysis On Time Scaling of Value-at-Risk Estimation

Quantitative Analysis - Econometrics

August 2023

CID: 01515252

Imperial College Business School

Client Specification

Our client is a significant unit within the Corporate & Investment Bank division of a globally renowned American financial institution, located in the United Kingdom. Their responsibilities include validating the development and use of models for market risk analysis, and reviewing firmwide portfolio risk perspectives to ensure the company's exposure remains within acceptable limits, which are subject to both US and UK regulations. As a guardian of financial risk management, they play a vital role in maintaining financial stability and safeguarding against potential market fluctuations for the firm.

In accordance with Basel II and Fundamental Review of the Trading Book, the bank is obligated to allocate assets to serve as a safeguard against large portfolio losses, often referred to 'Minimum Capital Requirements for Market Risk'. The bank is asked to compute the Value-at-Risk (hereafter referred to as VaR) on a daily basis using a 99th percentile. As part of the computation process, the bank is required to consider 'an instantaneous price shock equivalent to a 10 day movement in prices is to be used' (2009:p.13)[5]. The rule allows banks to calculate 10-day VaR directly, or alternatively, to scale a 1-day estimate to a 10-day horizon by applying an appropriate scaling factor. The client has adopted the latter approach for firmwide capital calculations.

A common practice for estimating the 10-day VaR involves calculating 1-day VaR and multiplying it by the square root of 10. In practice, known as the 'square-root-of-time-rule', is recommended by banking supervisors and extensively used the financial industry. However, this method has been criticized for systemic underestimating risk when applied to real data. This flaw is detrimental as it allows the portfolio managers to take on excessive risks. To address this issue, our client purposes an alternative method, called 'overlapping data aggregation' method, and seeks to explore the feasibility of its implementation.

Our role, as a consultant of the team, is to assess the client's existing approach by performing statistical tests on various null and alternative hypotheses. We will then estimate the potential gain or losses incurred when using the current method on their portfolios. Finally, we will provide an enhanced procedure, based on our findings and research, detailing how our client can calculate and report the 10-day VaR as required.

Note: Although the empirical analysis is based on the client's firmwide portfolio, due to confidentiality considerations, this project will utilise public available data in the empirical analysis section for illustration purpose.

Contents

Word Count of main body: 3256

1 Introduction

A standard method for estimating quantiles for a long period (in particular VaR) is by estimating 1-day VaR and then multiplying it by the square root of the time horizon (usually refers to the square-root-of-time rule in practice). This SRTR is recommended by banking supervisors and is widely used across the financial industry. However, such an approach is theoretically valid only if the daily returns are independent Gaussian random variables with zero mean, otherwise an approximation, and often leads to underestimation. Therefore, other approaches for estimating risk quantiles are proposed, one example being the use of data aggregation. In this project, the concern is on overlapping method: when aggregating the daily returns, the aggregation periods are allowed to overlap.

To compare the time-scaling approach and the overlapping data aggregation approach, a test ratio is proposed: $TestRatio = \frac{VaR_{ov}^{10d}}{\sqrt{10}VaR^{1d}}$, and its critical value is set to be 1. Our client adopts the following guidance: when the test ratio exceeds 1, further investigation will be conducted. However, this critical value is based on various assumptions that might not hold in realisation, hence it might not be a reliable signal. The aim of this project is to investigate this critical value and the approach using both simulation and real data in equity index, currency, and commodity price, and quantify the uncertainty associated with this choice of critical value. In addition, the project will give suggestions on the possible direction of modification, to avoid over- and/or under-estimation of 10-day VaR.

This project report is organised in the following way: the second section will use the Monte Carlo method to simulate returns data, and examine the impact of the documented stylised facts of asset returns to the test ratio. The third section will investigate the properties of historical data in 4 different asset classes and compare the gain/loss for using the time-scaling and overlapping approach. Two modified approaches to estimate 10-day VaR are then proposed in the fourth section, then all methods in the scope of this project are back-tested to compare their performance. Finally, the author will make some concluding remarks for the client.

Note: To make the solutions applicable in the business setting, throughout the project, the time horizon we consider for data aggregation is the 10-day holding period, and the length of data used for VaR calculation will be one year i.e. 264 trading days.

2 Simulation Study

In this section, two problems will be investigated. Firstly, we examine the use of critical value 1. As many financial models assume the returns are independently and identically distributed (i.i.d.) and they follow a Gaussian distribution, this will be our starting point. The problem is formulated as follows: Under the assumption that asset returns follow an i.i.d. Normal distribution with zero mean i.e. $X_t \sim \mathcal{N}(0,1)$, to what extent is the Test Ratio = 1 appropriate.

Next, as the assumption of i.i.d. Gaussian is not always true, we also explore different circumstances in which the asset returns might follow other specifications, and how the hypothesis tests will be affected.

To answer these two questions, we will first introduce the relevant concepts of quantile estimation, time scaling, data aggregation, the Monte Carlo technique, and hypothesis testing. Then we simulate returns, estimate 10-day VaR and analyse the results using the methods mentioned.

2.1 Methodology

2.1.1 VaR, SRTR and time aggregation

A portfolio's VaR at α confidence level with loss L is given by the smallest number l such that the probability that the loss L exceeds l is no larger than $1-\alpha$. Formally, as in [7], we define:

$$VaR_\alpha(L) = \inf\{l \in \mathbb{R} \mid P(L > l) \leq 1-\alpha\} = \inf\{l \in \mathbb{R} \mid F_L(l) \geq \alpha\},$$

where F is the distribution of the portfolio's profit and loss distribution. To align with the industry, according to Campell(2005) [2] we simplify the above definition and calculate

$$VaR_t(\alpha) = -F^{-1}(\alpha \mid \Omega_t),$$

where $F^{-1}(\cdot \mid \Omega_t)$ is the quantile function of the P&L loss distribution, which varies over time as market conditions and portfolio's composition, and with $\alpha = 99\%$.

From daily observation of return data, there are two main ways one can forecast long-period VaR. The first approach is to simply multiply the square root of 10 by the 1-day VaR calculation: $\sqrt{10}VaR^{1d}$.

The second approach is via time aggregation. Ten daily returns are first aggregated to 10-day returns, then from the aggregated returns, 10-day VaR is calculated. The aggregation can be done in two ways: non-overlapping, summing the returns in non-overlapping blocks; and overlapping, where the summing blocks are allowed to overlap.

The use of non-overlapping approach is favoured because it doesn't create dependency between data(Alexander, 2008)[1]. However, in the 264-day sample period, non-overlapping approach would pass only 26 observations to estimate the quantile. Therefore, the overlapping approach is more practical to use and is implemented in this project.

2.1.2 Monte Carlo Simulation

Assuming the data generating process is known, we generate samples of 100,000 realizations. To calculate 1d-VaR and overlapping VaR, for each realisation, the following steps are followed:

1. Generate 264 observations under different distributions X_i.
2. For the purpose of estimating overlapping VaR, an additional 9 observations are generated and make it a total of 273 observations X_i, then the aggreagted returns are $Z_j = \sum_{i=j}^{i+9} X_i$.
3. The two resulting sequences of data should both have 264 entries. We then calculate the VaR_{ov}^{10d} and VaR using quantile function in R.

The procedure is illustrated as the graph:

2.1.3 Hypothesis Testing and Power function

So far in this project, the returns we generated are assumed to follow an i.i.d Gaussian with zero mean. Given the well-documented stylised facts of asset returns, it is worth checking how well the 95% confi-

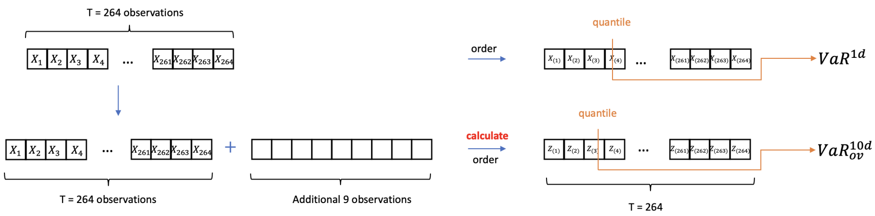


Figure 1: Monte Carlo Procedure

dence interval captures the risk of over/underestimation when the assumption is loosened. We will use the power of hypothesis testing to measure this risk, as defined below:

$$\beta(\alpha) = \text{P}(H_0 \text{ is not rejected} |\ H_1 \text{ is correct})$$
$$\text{and power} = 1 - \beta(\alpha),$$

where α = P(H_0 is not rejected| H_1 is correct) = Type I error, which is defined by our significance level.

A higher level of power indicates a lower beta i.e. Type II error, which means that the hypothesis test is better at detecting a false null hypothesis.

We use the Monte Carlo technique to generate the returns, in the data generation process, the following stylized facts are considered as our alternative hypotheses for the power tests:
Non-zero mean:

$$x_t = \mu + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,1) \tag{1}$$

Serial dependence using AR(1):

$$x_t = \rho x_{t-1} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,1) \tag{2}$$

Volatility clustering using GARCH(1,1):

$$x_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \omega + \alpha \sigma_{t-1}^2 + \beta x_{t-1}^2, \quad \epsilon_t \sim \mathcal{N}(0,1) \tag{3}$$

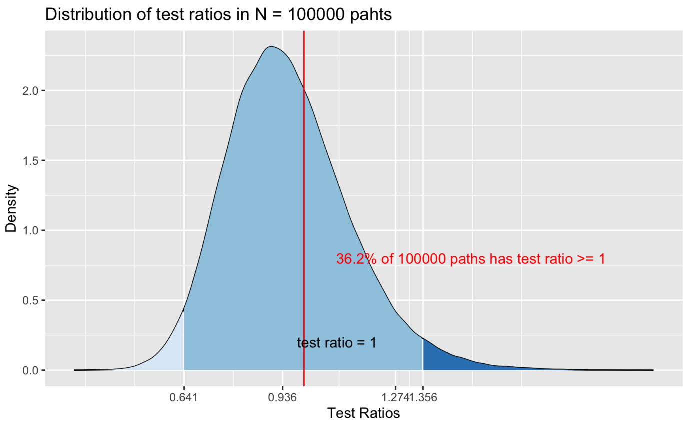
Heavy-tailedness using student-t distribution:

$$x_t \sim t(\nu) \tag{4}$$

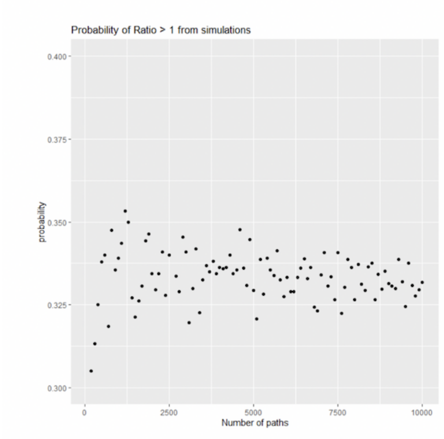
2.2 Results

In the first part, we answer the first question about the critical value 1 of the test ratio.

We first generate 100,000 paths of observations that follow i.i.d. Gaussian with zero mean using the Monte Carlo method and calculate the the test ratio for each path. The distribution of the test ratios for the 100,000 paths is shown in Figure 2a. In the figure, 36.2% of the time the ratio will be larger than one, and the wide 95% confidence interval [0.641, 1.356] indicates large simulation noise and a large number of false positives.



(a) Distribution of test ratios in N = 100000 samples



(b) Percentage of Test Ratio >=1 for different samples

Additionally, as Figure 2b shows, this simulation noise does not decrease monotonically as the number of generated paths increases and does not seem to converge to a single value. The probability Pratio >1 fluctuates around 31.5% - 36.5% for a large number of paths. This is to say, large simulation noise is inevitable.

Consequently, the answer to our first question is clear: given the large simulation noise, it would not be ideal to use the critical value 1. If such a test ratio must be implemented, an acceptable range, e.g. the 95% confidence interval, is more appropriate.

Next, we examine the following specific data generating processes with certain parameters using the Monte Carlo technique and power tests to answer our second question. The results are summarised in Table 1.

2.2.1 Non-zero mean

In the first scenario we assume returns are iid, drawn from Gaussian (μ, 1) as in 1 we let μ= -0.2,-0.15,-0.1, -0.05,0.05,0.1,0.15,0.2.

When the unconditional mean is negative, the overlapping approach tends to produce more conservative estimates. As the mean becomes positive, the ratio falls below 1, indicating an underestimation of the overlapping approach, compared to SRTR scaling.

The power under this alternative hypothesis is fairly low for negative mean returns, but it increases as the positive returns increase. However, even the highest power (0.32 when μ=0.2) in our experiment is not satisfactory, as it implies Type II error = 0.68.

2.2.2 Autocorrelated series AR(1)

To consider serial correlation, we draw samples from AR(1) as in 2, with ρ= -0.5,-0.4,-0.3,-0.2,-0.1,0.1,0.2, 0.3,0.4,0.5. Any other autocorrelation coefficients beyond or below these boundary levels are considered very less likely to happen, thus are not included in the study.

	Ratio	95%conf	Approximation Error (%)	Power
Panel A: Non Zero mean models with different means				
-0.2	1.1173	(0.821,1.498)	11.7342	0.0882
-0.15	1.0788	(0.778,1.466)	7.8843	0.0663
-0.1	1.0388	(0.742,1.428)	3.8810	0.0537
-0.05	0.9935	(0.686,1.390)	-0.6475	0.0441
0.05	0.8961	(0.582,1.304)	-10.3862	0.0784
0.1	0.8503	(0.533,1.271)	-14.9661	0.1297
0.15	0.7972	(0.477,1.219)	-20.2736	0.211
0.2	0.7463	(0.412,1.186)	-25.3669	0.32
Panel B: AR(1) models with different autocorrelation coefficients				
-0.5	0.5929	(0.408,0.840)	-40.7141	0.7076
-0.4	0.6571	(0.448, 0.925)	-34.2933	0.4883
-0.3	0.7225	(0.492,1.036)	-27.7453	0.2907
-0.2	0.7934	(0.537,1.121)	-20.6559	0.1466
-0.1	0.9478	(0.585,1.236)	-13.2906	0.0736
0.1	1.0339	(0.692,1.474)	3.3942	0.0746
0.2	1.129	(0.760,1.582)	12.8990	0.1415
0.3	1.2379	(0.873,1.749)	23.7877	0.2844
0.4	1.3624	(0.923,1.906)	36.2401	0.4825
0.5	1.5042	(1.008,2.076)	50.4233	0.6942
Panel C: GARCH(1,1) models with different alpha, beta parameters				
(0.130, 0.820)	0.9576	(0.578,1.502)	-4.2364	0.1216
(0.150, 0.800)	0.9546	(0.567,1.500)	-4.5387	0.1298
(0.130, 0.840)	0.9583	(0.563,1.504)	-4.4174	0.1346
(0.150, 0.820)	0.9568	(0.551,1.515)	-4.3243	0.1429
Panel D: Student-t models with different degrees of freedom				
2.5	0.8405	(0.470, 1.777)	-15.9464	0.3199
3	0.8256	(0.482,1.490)	-17.4430	0.2613
3.5	0.8224	(0.507,1.375)	-17.7619	0.2151
4	0.8259	(0.508,1.310)	-17.4119	0.1894
4.5	0.8342	(0.526,1.302)	-16.5783	0.1553
5	0.8362	(0.538,1.268)	-16.3774	0.1419

Note: Approximation Error = $\frac{VaR_{ov}^{10d} - \sqrt{10}VaR^{1d}}{\sqrt{10}VaR^{1d}} = Test\ Ratio - 1$, represents the percentage of overestimation if we assume the SRTR 10-day VaR is the "true" Var. If it's a negative value, then it implies an underestimation.

Table 1: Simulation results under different scenarios

We can see clearly from Table 1 Panel B 1 that, the overlapping approach will underestimate 10-day VaR when the returns are negatively autocorrelated, and will overestimate when the returns are positively autocorrelated. In terms of power, regardless of the sign, the more autocorrelated a return series is, the higher the power of the test.

2.2.3 Volatility Clustering

To incorporate volatility clustering characteristic of asset returns, we simulate GARCH(1,1) model as in 3, experimenting with parameters $\omega = 0.0001, \alpha + \beta = 0.95$ or 0.97, the choice of parameters is inspired by Wang et al (Wang, Yeh & Cheng, 2011)[9]. For all 4 cases, the powers are very low. When examining the ratios, interestingly, the use of the overlapping approach underestimates the risk more than using SRTR approach. However, since the approximation error is below 5%, arguably the estimates are fairly similar and the ratio bias is small. As the parameters chosen are fairly similar, it might be worth considering other combinations as well. In the next section, we fit a GARCH model and use the fitted model to repeat the test.

2.2.4 Heavy-Tailedness using student-t distribution

In the fourth scenario we assume returns are i.i.d., drawn from student-t distribution as in 4 with 2.5 – 5 degrees of freedom, to see the impact of fatter tails. The choice of parameters aligns with the most financial literature.

The test ratio is below 1 in all six cases, meaning that on average the square-root-of-time is the more conservative estimate. Therefore, in the case of heavy-tailedness, switching to the overlapping method will be problematic.

The powers of hypothesis test are still very low, and it decreases as the distribution becomes less heavy-tailed. However, the overall power under the heavy-tailedness alternative hypothesis is higher than any other alternative hypotheses considered in the project.

2.2.5 Summary

To conclude and answer the second questions, the approximation errors tell us that serial dependence and heavy-tailedness are significant factors that might impact the performance of the scaling of VaR, and the powers indicate that our test ratio and hypothesis test are not appropriate to detect between the assumed i.i.d. Normal returns and the returns normally seen in the real-world context. However, the parameters chosen in this section is arbitrary, as we solely want to isolate the different effects of different data generating processes. In the next section, we further conduct empirical study to examine the impact in a real business setting,

3 Empirical Study

In this section, we examine the characteristics of asset returns based on real data and test the applicability of our test ratio and hypothesis tests. We aim to answer the two questions in the previous section under the real market conditions.

3.1 Data

Using publicly available data source from Investing.com [6], we choose one representative product from each asset classes: SP500 - indices, USD/BRP - exchange rate, BMW - equity share, and Gold Futures – commodity future. All currencies are converted to USD. Two periods are chosen to represent normal market conditions and market under stressed: 21st July 2022 – 25th July 2023, and 22nd November 2007 – 25th November 2008, with 273 returns in each period.

3.2 Characteristics of returns

	2007 - 2008				2022-2023			
	SP500	BMW	USD/GBP	Gold Futures	SP500	BMW	USD/GBP	Gold Futures
Mean	-0.0018	-0.0024	0.0011	0.0003	0.0007	0.0016	-0.0003	0.0006
AR Coefficient	-0.1466	-0.0684	0.0999	0.1271	-0.0056	-0.0618	0.0684	-0.0441
Tail Index	1.946	3.095	2.875	2.3567	3.248	2.990	2.739	3.004
96% Conf Tail Index	(0.888,3.004)	(1.412,4.778)	(1.312,4.438)	(1.076, 3.638)	(1.482,5.013)	(1.365,4.615)	(1.250,4.227)	(1.371,4.636)

Table 2: Estimated parameters

For the chosen historic data, we summarised some features of the returns of each asset in the two periods. The features are corresponded to the 4 alternatives hypotheses in the previous section, including mean, autocorrelation coefficient, tail index [1] and GARCH(1,1) parameters.

	μ(e-04)	ω(e-06)	α	β
SP500 08	-7.875	9.735	0.1005^{***}	0.8644^{***}
SP500 23	12.304	0.946	0.0514^{**}	0.9477^{***}
BMW 08	-20.474	17.5	0.0863^{**}	0.8887^{***}
BMW 23	14.049	14.497	0.016	0.9301^{***}
USD/GBP 08	6.813	1.32	0.0925^{**}	0.8784^{***}
USD/GBP 23	-1.958	1.516	0.0915^{**}	0.8806^{***}
Gold Futures 08	6.571^{*}	9.257^{*}	0.0766	0.8911^{***}
Gold Futures 23	3.901	6.171	0.0323	0.8957^{***}

Note: the "*" mark represents the level of statistical significance, "*" - the significance level is less than or equal to 0.1; "**" - the significance level is less than 0.01; "***" - the significance level is less than 0.001

Table 3: Estimated GARCH(1,1) model parameters

Some conclusions can be drawn from the results in Table 2 and 3:

1. Market is efficient: Except USD/BRP exchange rate in 2007 – 2008, none of the means is significantly different from 0.

2. No strong dependence is witnessed: The autocorrelation coefficients are not significant and are limited within (-0.15, 0.15).

3. The assets are all heavy-tailed, in both normal and stressed market conditions: Apart from SP500 in 2007 – 2008, the lower boundary of the 95% confidence interval of the estimated tail index of all other assets are greater than 1.

4. Volatility clustering exists: Most estimated alpha and beta parameters are statistically significant at 0.05 level, and $\hat{\alpha} + \hat{\beta} < 1$. However, interestingly, the parameters are quite different from what we assumed in the simulation part, which means we need further calibration of the model.

[1] estimated using log-log rank-size regression method with 0.5 adjustment on the rank, as suggested by Ibragimov et al(Ibragimov, Ibragimov & Walden)[4]

3.3 Results

Combining the simulation results and empirical results in Section 2 and 3, only the highlighted rows in Table 1 fall under our concern. By comparing the approximation errors, we conclude that in most cases, the error is negative, meaning that for our data, the overlapping approach will lead to underestimation. Among all relevant cases, Panel D in Table 1 shows that overlapping method yields the most severe level of underestimation when the returns exhibit heavy-tails. One possible reason is that, under heavy-tailedness, the use of overlapping method is not ideal because it is smoothed out by aggregating daily returns.

In conclusion, heavy-tailedness in returns dominates among all the other factors within the scope of this report. Although volatility clustering is witnessed, it does not generate severe approximation error as heavy-tailedness does.

4 Modification and Backtesting Results

Two insights can be deduced from previous two sections:1) our test ratio and hypothesis tests fail to detect changing return characters. 2) for a more precise estimate of 10-day VaR, we need to incorporate heavy-tailedness and volatility clustering. Inspired by literature, this section will introduce two modifications on the SRTR approach, to increase the level of precision. At last, we will backtest the above methods on four historical series in Section 3.1, with an extended period from January 2007 to July 2023.

4.1 Tail-Adjusted SRTR VaR

After conducting a simulation study, Wang et al (Wang, Yeh & Cheng, 2011)[9] found that weak dependence in returns dominates among all the stylised facts, thus they propose adjusting the SRTR scale with a proxy for serial dependence: $MVaR(h) = \sqrt{10 * VR}VaR^{1d}$, where VR standards for Variance Ratio, a test statistics often employed for testing market efficiency.

By implementing a similar idea, we propose adjusting the SRTR VaR by multiplying a proxy of heavy-tailedness, as we found that heavy-tailedness is the most significant factor. We choose a moderate measure: the lower endpoint of a 95% confidence interval of a Hill estimate, and it serves as a pretest of the applicability of SRTR. Thus, our estimate of 10-day VaR becomes: $VaR^{10d}_{Tail} = \sqrt{10\zeta^{\hat{Hill}}_{0.95}\mathbb{1}_{\hat{\zeta}\geq 1}}VaR^{1d}$

We choose this empirical adjusting factor because it can reflect recent market conditions. Since we only have 264 observations, the estimated index is subject to uncertainty, thus we use the endpoint of the confidence interval, instead of the estimator itself.

4.2 EVT-based multi-day VaR

This approach is proposed by McNeil et al(McNeil & Frey, 2000)[8]. We think it is suitable to explore this method in this project because we witness volatility clustering in our historical data. In addition, given the fact that we have small sample size for prediction (since we only use 264 observations for each VaR computation), the bootstrapping component embedded in this method is appealing to increase precision.

The basic idea of this approach is summarised as follows:

1. In a lookback period of 264 days, fit an AR(1) – GARCH(1,1) model to the negative log return series using a pseudo-maximum-likelihood approach, calculate the residuals.

2. Apply Extreme Value Theory to model the tail of the residuals.
3. Bootstrap from the residuals, if the selected residual exceeds certain thresholds, then sample from the fitted tail model as an excess to the residual. This step gives a noise distribution.
4. Use the composite estimated noise distribution and the fitted GARCH-type model to simulate 1000 paths of future 10 days returns, and calculate the corresponding sums of the 10 returns. These 1000 sums follow the distribution $F_{X_{t+1}++X_{t+10}}$.
5. Finally, the EVT-baesd 10-day 99% VaR is defined as $VaR^{10d}_{EVT} = \inf\{x \mid F_{X_{t+1}+...+X_{t+10}} \geq 0.99\}$.

4.3 Backtest and Results Analysis

We use the daily negative log returns of SP500, BMW, USD/GBP exchange rate and Gold Futures from January 2007 to July 2023 (about 4000 observations) to backtest the approaches of estimating 10-day VaR: SRTR time-scaling approach, overlapping approach, tail-adjusted SRTR approach, and EVT-based multi-day approach.

To backtest on a historical series containing N (N >4000) observations, we calculate the 10-day VaR on t-th ($t \in \{273, ..., N-10\}$) using a window of 264 days. For methods involving estimating tails, a truncation rate of 0.1 is fixed. Then, we compare each estimate with the cumulative sum $x_{t+1} + ... + x_{t+10}$. If our estimated 10-day VaR is less than the cumulative sum, then a breach is said to occur. Given the series length N and quantile 0.99, $0.01 * N$ number of breaches are expected to occur, which is set as our benchmark. In the table below, for each estimating method and each asset, the number of breaches is presented.

	SP500	BMW	USD/GBP	Gold Futures
length of test	4168	4053	4321	4277
Expectation	42	41	43	43
SRTR	60	62	39	43
Overlapping	112	108	96	79
Tail-Adjusted	48	37	22	33
EVT-based	32	6	34	20

Table 4: Backtesting results

From Table 4, SRTR clearly underestimates 10-day VaR of SP500 and BMW stock, however, it performs well on exchange rate and gold futures. The overlapping 10-day VaR underestimates in all 4 assets, which aligns with the result in simulation study. The tail-adjusted SRTR VaR corrects the underestimation on index and stock, but it leads to slight overestimations for the other two assets. This suggests for exchange rate and gold price, SRTR might be the best approach. The same argument is mentioned by McNeil et al(McNeil & Frey, 2000)[8] as well.

EVT-based VaR gives the most conservative estimations. We realise this might be a consequence of using short horizon to estimate GARCH parameters. Additionally, this method takes long time to run and the output is not stable: 1000 paths are not enough for the quantile to converge. In the literature, the authors apply a second round of EVT on the simulated observations to increase precision. Due to concern in computational time, we did not do the same. We recognise this as the reason of unstable output and overestimation.

5 Conclusion

The project does not recommend continuing to use the test ratio = 1 as a signal for further investigation, because it absorbs large uncertainties, and does not react to changes in the characteristics of returns. The author understands the simplicity of this method, thus if the client insists on using such test ratio, the 95% confidence interval should be used with care instead.

Alternatively, the author recommends two modifications to the existing methods based on the discovery of the significance of heavy-tailedness and volatility clustering. As empirical evidence suggests, these two methods are more stable and yield more accurate estimates if one wants to avoid risk underestimation. Nevertheless, the two methods are empirically derived hence lack examination of robustness and theoretical support, further steps should be taken to ensure the soundness of the potential use of such approaches.

To take the project further, one can consider including the impact of jumps as Danielsson et al (Danielsson & Zigrand, 2003)[3] and Wang et al (Wang, Yeh & Cheng, 2011)[9] did. On the other hand, since the estimates are derived from only 264 data points, some techniques developed for small sample size are good next steps to be considered.

6 References

[1] Carol Alexander. *Market risk analysis. Volume IV, Value-at-risk-models.* 1st Edition. Wiley, 2008.

[2] Sean D. Campbell. "A Review of Backtesting and Backtesting Procedures". In: *Finance and Economics Discussion Series* 2005 (May 2005), pp. 1–23. DOI: `10.17016/feds.2005.21`.

[3] Jón Daníelsson and Jean-Pierre Zigrand. *On Time-scaling of Risk and the Square-root-of-time-rule.* Financial Markets Group Discussion Papers (439). Financial Markets Group, The London School of Economics and Political Science, London, UK., 2003. URL: `http://eprints.lse.ac.uk/id/eprint/24827`.

[4] Marat Ibragimov, Rustam Ibragimov, and Johan Walden. *Heavy-tailed distributions and robustness in economics and finance.* Springer Verlag, 2015.

[5] Bank for International Settlements. *Basel Committee on Banking Supervision Revisions to the Basel II market risk framework.* 2009. URL: `https://www.bis.org/publ/bcbs158.pdf`.

[6] Investing.com. *Investing.com - Stock Market Quotes Financial News.* Investing.com, 2023. URL: `https://www.investing.com/`.

[7] Alexander J Mcneil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management : concepts, techniques and tools.* Princeton University Press, 2015.

[8] Alexander J. McNeil and Rüdiger Frey. "Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach". In: *Journal of Empirical Finance* 7 (Nov. 2000), pp. 271–300. DOI: `10.1016/S0927-5398(00)00012-8`. URL: `https://www.sciencedirect.com/science/article/abs/pii/S0927539800000128`.

[9] Jying-Nan Wang, Jin-Huei Yeh, and Nick Ying-Pin Cheng. "How accurate is the square-root-of-time rule in scaling tail risk: A global study". In: *Journal of Banking Finance* 35 (May 2011), pp. 1158–1169. DOI: `10.1016/j.jbankfin.2010.09.028`.

7 Appendix

7.1 Code

For reproducibility, the data and the R notebooks containing the code used in this project are available at `https://drive.google.com/drive/folders/163DPW8QOqVEIAyVZEJn8vWbouxNC_bJU?usp=share_link`

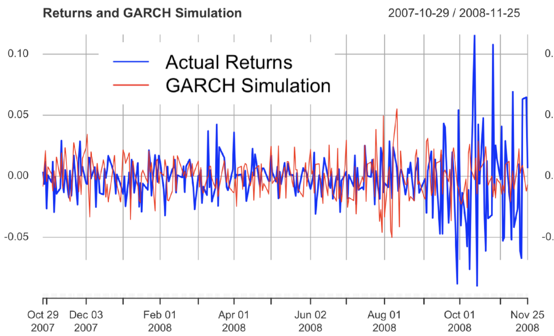
Most of R packages used are documented in the notebooks. One package "gPdtest" was used to perform simulation and estimation of Generalised Pareto Distribution of the following form:

$$G_{\zeta,\beta}(y) = \begin{cases} 1-(1+\zeta y/\beta)^{-\frac{1}{\zeta}}, & \text{if } \zeta \neq 0, \\ 1-exp(-y/\zeta), & \text{if } \zeta = 0. \end{cases} \tag{5}$$

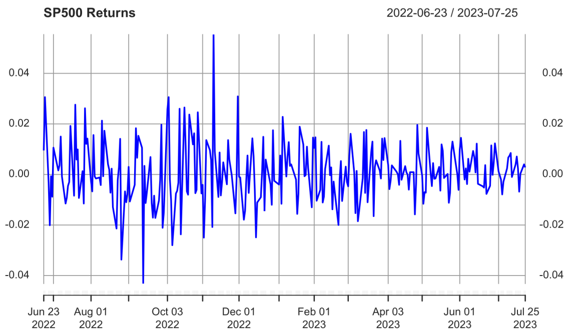
where β >0 and the support is $y \geq 0$ when $\zeta \geq 0$ and $0 \leq y \leq -\beta/\zeta$ when $\zeta < 0$.
It has been archived and the functions used in this project are available at: `https://cran.r-project.org/web/packages/gPdtest/index.html` in "gPdtest$_0$.4.$tar.gz$".

7.2 Graphs of asset returns

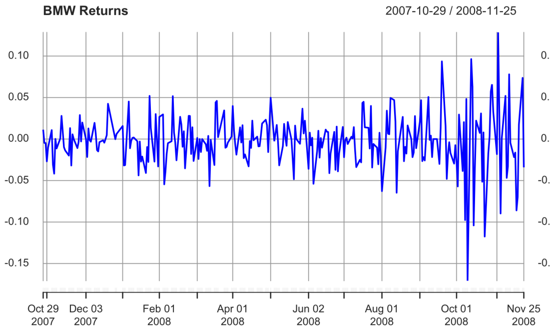
Attached are some graphs of the asset returns in period 2007 - 2008, and 2022 - 2023.



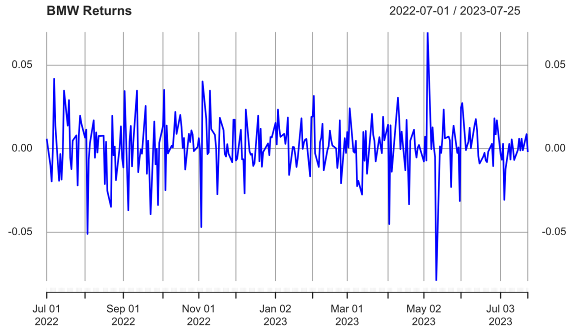
(a) SP500 2007-2008 returns and fitted GARCH(1,1) values



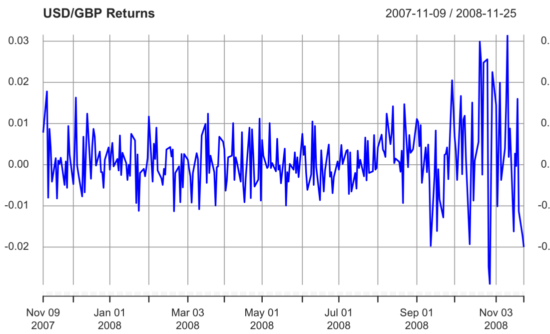
(b) SP500 2022-2023 returns



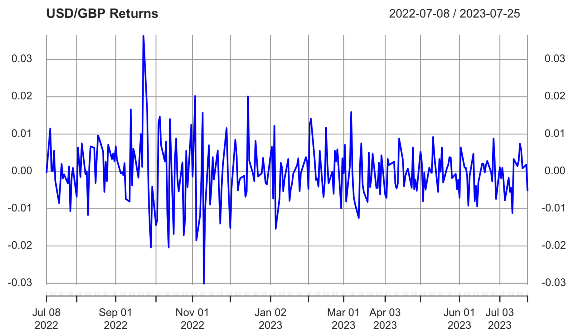
(a) BMW Stock 2007-2008 returns



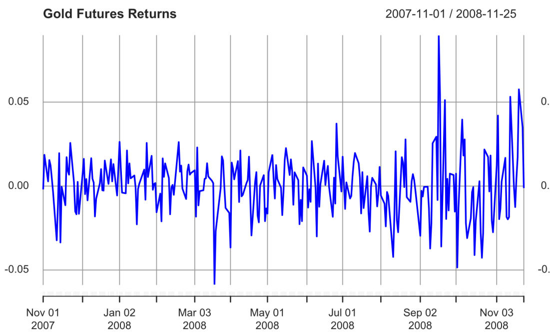
(b) BMW Stock 2022-2023 returns



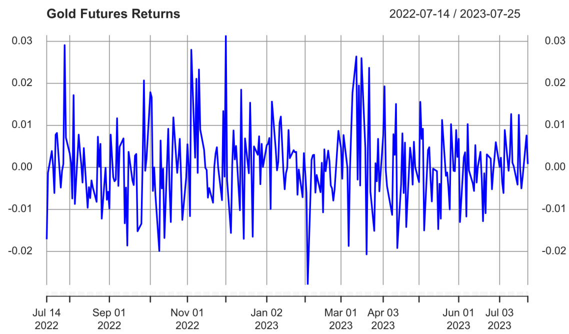
(a) USD/GBP Exchange Rate 2007-2008 returns



(b) USD/GBP Exchange Rate 2022-2023 returns



(a) Gold Futures 2007-2008 returns



(b) Gold Futures 2022-2023 returns