Empirical Methods in Finance

Project #2: Dynamic Allocation and VaR of a Portfolio

Group 09:
Ivan Khalin
Litian Zhang

PART 1 – Static Allocation

Q1.1

The optimal weights are found by taking the first order derivative of the mean-variance criterion and setting it to 0, which yield the following expression:

$$\max_{\{\tilde{\alpha}\}} \mu_p - \frac{\lambda}{2}\sigma_p^2 \quad \Longrightarrow \quad \max_{\{\tilde{\alpha}\}} \tilde{\alpha}'\mu + (1 - e'\tilde{\alpha})R_f \; - \frac{\lambda}{2}\tilde{\alpha}'\Sigma\tilde{\alpha}$$

$$\text{FOC: } \mu - e'R_f - \Sigma\tilde{\alpha} = 0 \quad \Longrightarrow \quad \tilde{\alpha}^* = \frac{1}{\lambda}\Sigma^{-1}(\mu - e'R_f)$$

$$\text{where: } \Sigma = \begin{bmatrix} \sigma_s^2 & \sigma_{sb} \\ \sigma_{sb} & \sigma_b^2 \end{bmatrix}$$

Thus, the individual weights are then given by:

$$\begin{bmatrix} \tilde{\alpha}_s \\ \tilde{\alpha}_b \end{bmatrix} = \frac{1}{\lambda}\frac{1}{\sigma_s^2\sigma_b^2 - \sigma_{sb}^2}\begin{bmatrix} \sigma_b^2 & -\sigma_{sb} \\ -\sigma_{sb} & \sigma_s^2 \end{bmatrix}\begin{bmatrix} \mu_s - R_f \\ \mu_b - R_f \end{bmatrix}$$

$$\Longrightarrow \begin{cases} \tilde{\alpha}_s = \dfrac{1}{\lambda}\dfrac{1}{\sigma_s^2\sigma_b^2 - \sigma_{sb}^2}\left[\sigma_b^2(\mu_s - R_f) - \sigma_{sb}(\mu_b - R_f)\right] \\ \tilde{\alpha}_b = \dfrac{1}{\lambda}\dfrac{1}{\sigma_s^2\sigma_b^2 - \sigma_{sb}^2}\left[\sigma_s^2(\mu_b - R_f) - \sigma_{sb}(\mu_s - R_f)\right] \end{cases}$$

Q1.2

Simply plugging in the numbers, we obtain the following weights for the static strategy:

λ	Stock	Bond	Risk-Free Asset
2	1.47	-0.35	-0.11
10	0.29	-0.07	0.78

It then becomes clear how the risk aversion parameter λ governs the weights of the two strategies. The weights for the stock and bond are always proportional to each other with $\tilde{\alpha}_{s,b}^{\lambda=2} = 5\tilde{\alpha}_{s,b}^{\lambda=10}$. This relation equally holds for the dynamic allocations, though importantly, the weight for the risk-free asset does not follow this relationship and must be computed anew every time.

Additionally, in estimating these optimal weights and in cross validating the results across team members, it became apparent that the optimal mean-variance weights are extremely sensitive to data inputs. For instance, one team-member's expected return was ever so slightly different, since they had opted to backfill the first observation for returns. Having just one additional value in a 1200 long sample resulted in a 3.6% and 1.7% change in the stock and bond weights respectively. Subsequently, we opted to simply remove the 'NAN' for first observation.

[Part 2 Below]

PART 2 – Estimation of a GARCH Model
Q2.1
The Table 2.1.1 presents solid evidence for the non-normality characteristic of both the stock and the bond. In a Kolmogorov-Smirnov test, the null hypothesis H_0 proposes that the data follows a normal distribution, opposing to the alternative hypothesis H_a that the data does not follow a normal distribution. We can see that for both the stock and the bond that H_0 is strongly rejected, given that their p-values are significantly smaller than 5%. Therefore, both excess returns and squared excess returns for the stock and the bond do not follow a normal distribution.

Table 2.1.1: Test Results for Non-normality (Kolmogorov-Smirnov Test)			
Excess Returns		**Test Statistic**	**P-value**
	Stocks (S&P500)	0.083178	1.136957e-8
	Bond (DAX)	0.075026	2.546712e-8
Squared Excess Returns			
	Stocks (S&P500)	0.358607	1.149728e-138
	Bond (DAX)	0.339205	1.056947e-123

Next, we implement a Ljung-Box Test with 4 lags to test for the auto-correlation for the bond and the stock. The null hypothesis H_0 that the data is not auto-correlated is tested against the alterative hypothesis H_a that the autocorrelation is present in the data. For the squared excess returns, the p-value presented in the Table 2.1.2 shows that we can safely reject the null hypothesis under 95% confidence interval, suggesting that the auto-correlation is not significant for both the bond and the stock within 4 lags. However, the opposite is true for the excess returns, with the null hypothesis not rejected under 95% confidence interval, for both stock and bond. Thus, these results suggest there is non-negligible degree of auto-correlation of excess returns within 4 lags.

Table 2.1.2: Test Results for Auto-Correlation (Ljung-Box Test, 4 lags)			
Excess Returns		**Test Statistic**	**P-value**
	Stocks (S&P500)	8.039479	0.090143
	Bond (DAX)	6.622373	0.157241
Squared Excess Returns			
	Stocks (S&P500)	361.846353	4.852335e-77
	Bond (DAX)	240.380883	7.681089e-51

Q2.2
The AR(1) model is defined as:

$$R_{t+1} = \alpha + \rho R_t + \varepsilon_{t+1} \quad \text{or} \quad \begin{bmatrix} R_{t+1} \\ \vdots \\ R_T \end{bmatrix} = \begin{bmatrix} 1 & R_t \\ & \vdots \\ 1 & R_{T-1} \end{bmatrix} \begin{bmatrix} \alpha \\ \rho \end{bmatrix} + \varepsilon$$

With coefficients computed as:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\rho} \end{bmatrix} = \left[\begin{bmatrix} 1 & R_t \\ & \vdots \\ 1 & R_{T-1} \end{bmatrix}' \begin{bmatrix} 1 & R_t \\ & \vdots \\ 1 & R_{T-1} \end{bmatrix} \right]^{-1} \begin{bmatrix} 1 & R_t \\ & \vdots \\ 1 & R_{T-1} \end{bmatrix}' \begin{bmatrix} R_{t+1} \\ \vdots \\ R_T \end{bmatrix}$$

And the residuals:

$$\hat{\varepsilon}_{t+1} = R_{t+1} - (\hat{\alpha} + \hat{\rho} R_t)$$

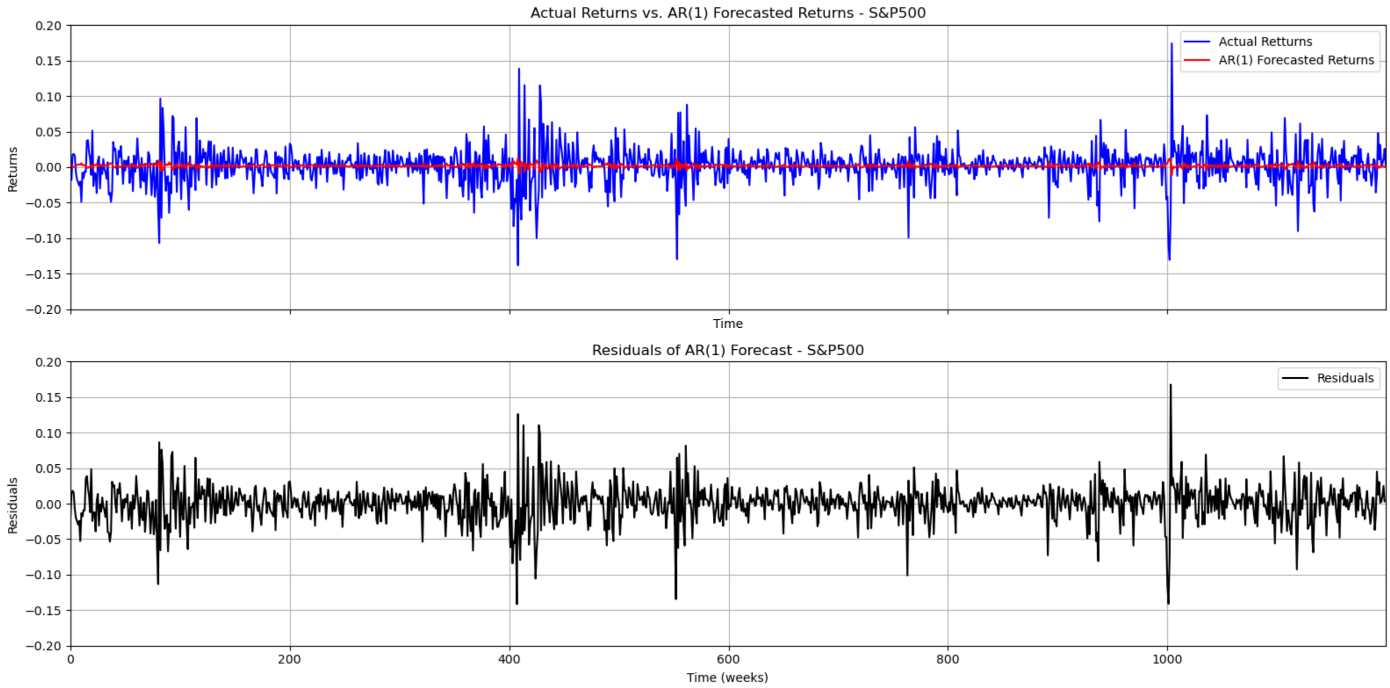
The parameter estimates are the following:

Figure 2.2.1: AR(1) model Statistics						
Stock (S&P500)		**Coefficient**	**Standard Error**	**P-value**	**AIC**	**BIC**
	Constant (α_s)	0.0019	0.001	0.010	-4645.271	-4630.003
	Slope (ρ_s)	-0.0769	0.029	0.008		
Bond (DAX)						
	Constant (α_b)	0.0016	0.001	0.120	-5369.249	-5353.982
	Slope (ρ_b)	-0.0151	0.029	0.601		

The parameter estimates for stock (α_s, ρ_s) are both statistically significant when evaluated at a 95% confidence interval, suggesting that the AR(1) model is a valid model for capturing the dynamics of returns for the stock. However, the p-value for bond parameter estimates is not significant when evaluated at the same confidence interval. This suggests the AR(1) might not be the best choice to model the return dynamics of the bond.

The AIC and BIC are the two methods we chose for assessing the goodness-of-fit of the model, the low values show generally good fit in terms of error minimization, with a penalty for the number of parameters used. Inspired by the statistical insignificance of the parameter estimates of the bond, we conjecture if an AR model with more lags can improve the significance level and ameliorate the goodness-of-fit. The AR(2) model was performed for the bond, but the result is rather disappointing. Both AIC and BIC have decreased by tiny amount ($AIC = -4640.773, BIC = -4620.419$), this suggests that adding complexity does not improve the goodness-of-fit of the model. The parameter estimates are still statistically insignificant with p-values being 0.584 and 0.124 for $\rho_b^{(Lag1)}$ and $\rho_b^{(Lag2)}$ respectively. We conclude that adding more lags to the AR model does not significantly improve the statistical accuracy. Therefore, we decide to tolerate that the parameter estimates for the bond are not statistically significant and proceed to the next steps.

Plotting the AR(1) model predictions for the stock market returns, it becomes apparent that the AR(1) model is not an adequate tool in forecasting the volatility dynamics of the stock. Particularly, it falls short in the periods of high volatility, which is captured by spikes in the AR(1) forecast residuals on the figure below. This lack of explanatory power in periods of high volatility is what motivates us to model conditional volatility.



Q2.3

The GARCH(1,1) model allows us to model the conditional volatility dynamics of a time series process. To estimate the parameters of the GARCH(1,1) model defined as:

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

We use numerical optimization to maximize the conditional log likelihood function:

$$\log L_T(\theta) = \sum_{t=1}^{T} \log\big(\ell_t(\theta)\big) \quad \text{where} \quad \log\big(\ell_t(\theta)\big) = \sum_{t=1}^{T} -\frac{1}{2}\left(\log(2\pi) + \log(\sigma_t^2) + \frac{\varepsilon_t^2}{\sigma_t^2}\right)$$

The unconditional variance is assumed to be $\bar{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_t$ with T the length of the vector of residuals from the AR(1) model, implying that the first conditional variance estimate is made for $t = 2$. The log likelihood maximization yields the following parameter estimates:

Figure 2.3.1: GARCH Estimates for $R_{i,t+1}$				
	ω	α	β	$\alpha+\beta$
Stock	0.000033	0.23104	0.729992	0.961030
Bond	0.000052	0.17174	0.791981	0.963720

ω measures the 'ambient' volatility, that is, the baseline volatility in the absence of innovations in the previous period. α could be thought of as measuring the sensitivity of volatility to shocks in the previous period and β the persistence of volatility. Together, $\alpha+\beta$ measure the holistic persistence of volatility, hence, a value larger than 1 would imply that the conditional variance is a non-stationary process that would converge to infinity, since the parameters ω, α, β are constrained to be positive. In theory, even if $\alpha+\beta>1$, the conditional variance process can still be strictly stationary i.f.f $\omega>0$ and $E[\log(\alpha_1 z_t^2+\beta_1)]<0$[1], due to the Jensen's inequality, we have:

$$E[\log(\alpha_1 z_t^2+\beta_1)]<\log(E[\alpha_1 z_t^2+\beta_1])=\log(\alpha_1+\beta_1) \quad \text{since} \quad z_t \sim iid(0,1)$$

Our parameter estimates suggest that the baseline volatility for the stock is lower than the bond. Both the stock and the bond markets show a highly persistent volatility with the sum of α and β being very similar. Notably, the stock market is more sensitive to innovations than the bond market, as suggested by the higher α.

Next, we test the null hypothesis that $\alpha+\beta=1$ for both the stock and the bond, with both a t and Wald test. We do so by determining the asymptotic variance of the GARCH parameters as the inverse of the Fisher information matrix, which for a GARCH(1,1) model is defined as:

$$\hat{I}(\theta)=\frac{1}{2T}\sum_{t=1}^{T}\frac{1}{\hat{\sigma}_t^4}\begin{pmatrix}1 & \hat{\varepsilon}_{t-1}^2 & \hat{\sigma}_{t-1}^2 \\ \hat{\varepsilon}_{t-1}^2 & \hat{\varepsilon}_{t-1}^4 & \hat{\varepsilon}_{t-1}^2\hat{\sigma}_{t-1}^2 \\ \hat{\sigma}_{t-1}^2 & \hat{\varepsilon}_{t-1}^2\hat{\sigma}_{t-1}^2 & \hat{\sigma}_{t-1}^4\end{pmatrix}$$

Then, the tests are defined as:

$$\begin{aligned} &\text{Wald test:} \quad T\frac{(\alpha+\beta-1)^2}{Var(\alpha+\beta)} \\ &t\text{ test:} \quad \sqrt{T}\frac{\alpha+\beta-1}{\sqrt{Var(\alpha+\beta)}} \end{aligned} \qquad \text{with} \quad Var(\alpha+\beta)=Var(\alpha)+Var(\beta)+2Cov(\alpha,\beta)$$

Where the variances and the covariance are taken from the inverse of the fisher information matrix:

$$\begin{aligned} Var(\alpha)&=\hat{I}(\theta)^{-1}_{\{2,2\}} \\ Var(\beta)&=\hat{I}(\theta)^{-1}_{\{3,3\}} \\ Cov(\alpha,\beta)&=\hat{I}(\theta)^{-1}_{\{2,3\}}=\hat{I}(\theta)^{-1}_{\{3,2\}} \end{aligned}$$

The test results are summarized in tables 2.3.2 and 2.3.3 below. Seeing as in this context of testing a single parameter restriction the Wald test is simply the t-test squared, the test statistics for the two procedures lead to identical conclusions. Both tests suggest a confident non-rejection of the null, the implications of which were discussed above. Our test results confirm the situation suggested by Nelson, and thus not raising concerns.

Table 2.3.2: GARCH Parameter Test at 95% Confidence Interval (*Wald Test*, *df* = 1)				
	Coefficient	**Standard Deviation**	**Test Statistics**	**Critical Value** ($\chi^2_{(1)}, \alpha=5\%$)
Stock	0.961030	2.675010	0.254472	3.841459
Bond	0.963720	2.860082	0.192930	3.841459

[1] Nelson, D. B. (1990). Stationarity and Persistence in the GARCH(1,1) Model. *Econometric Theory*, *6*(3), 318–334.

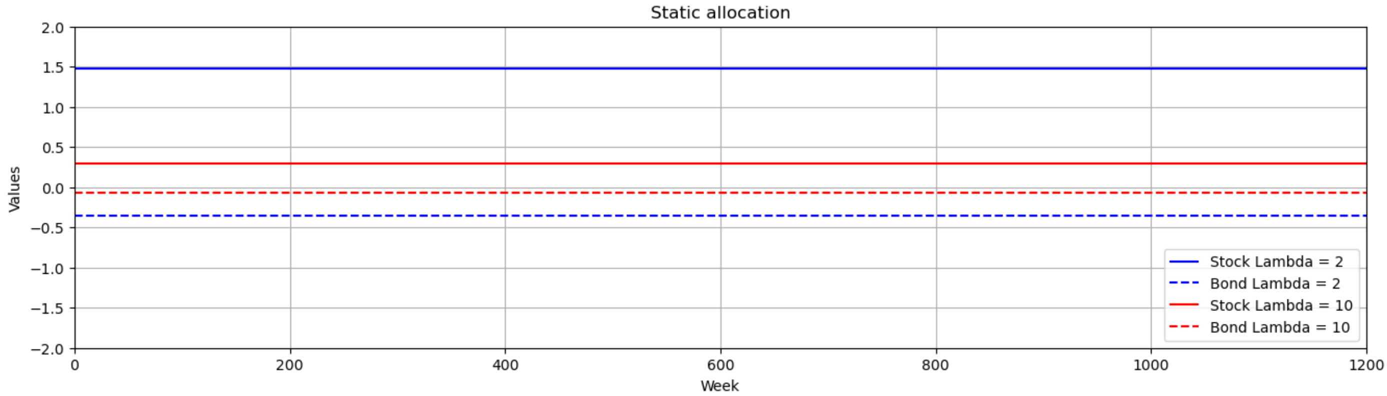
Table 2.3.3: GARCH Parameter Test at 95% Confidence Interval (T-*test*)				
	Coefficient	**Standard Deviation**	**Test Statistics**	**Critical Value** $(t, \alpha = 5\%)$
Stock	0.961030	2.675010	-0.504452	1.644854
Bond	0.963720	2.860083	-0.439238	1.644854

[Part 3 Below]

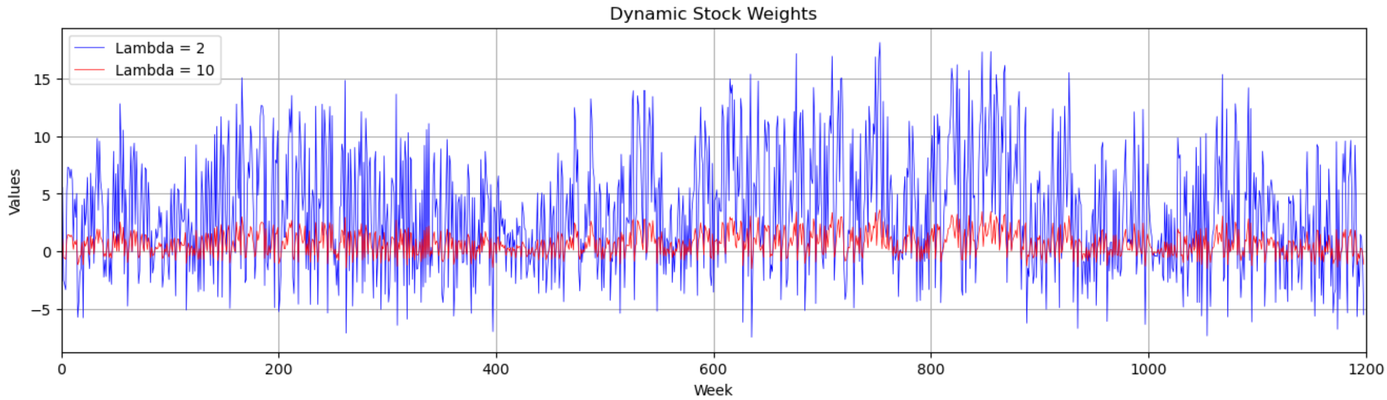
PART 3 – Dynamic Allocation

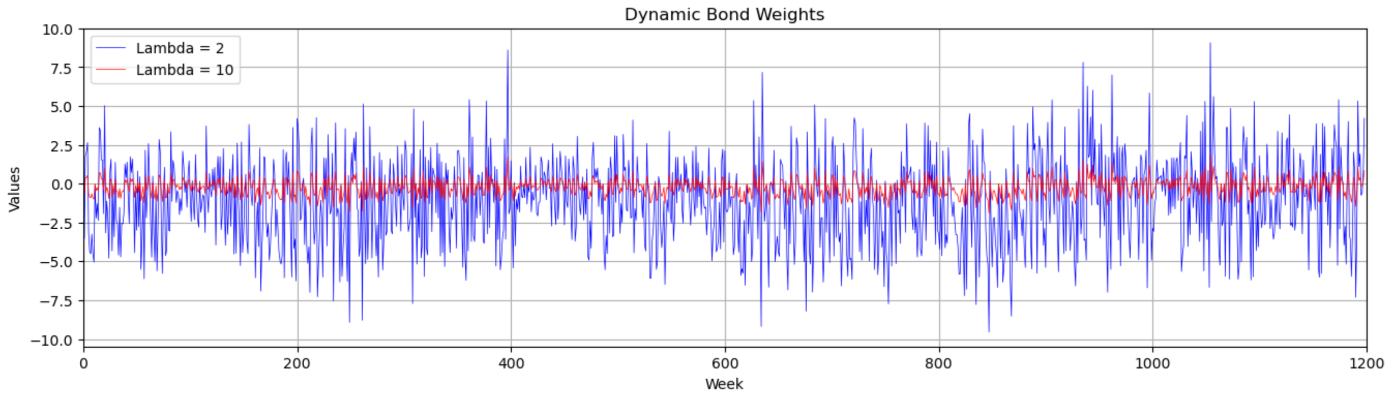
Q3.1

The method for estimating the optimal weights for the dynamic strategy closely resembles the procedure described in Part 1, with the difference being the use of dynamic expected returns coming from the AR(1) model, dynamic covariance matrix coming from the GARCH(1,1) model, and the use of a non-constant risk-free rate. The static optimal weights $\tilde{\alpha}^*$ for the strategy with constant expected returns and volatility do not change throughout the sample and are represented by straight lines:



The dynamic optimal weights $\tilde{\alpha}_t^*$ for the strategy with conditional expected returns and variance are highly volatile, and often extreme. For the $\lambda = 2$ strategy in particular, weights can reach massive short positions, approaching 10 times the total wealth:



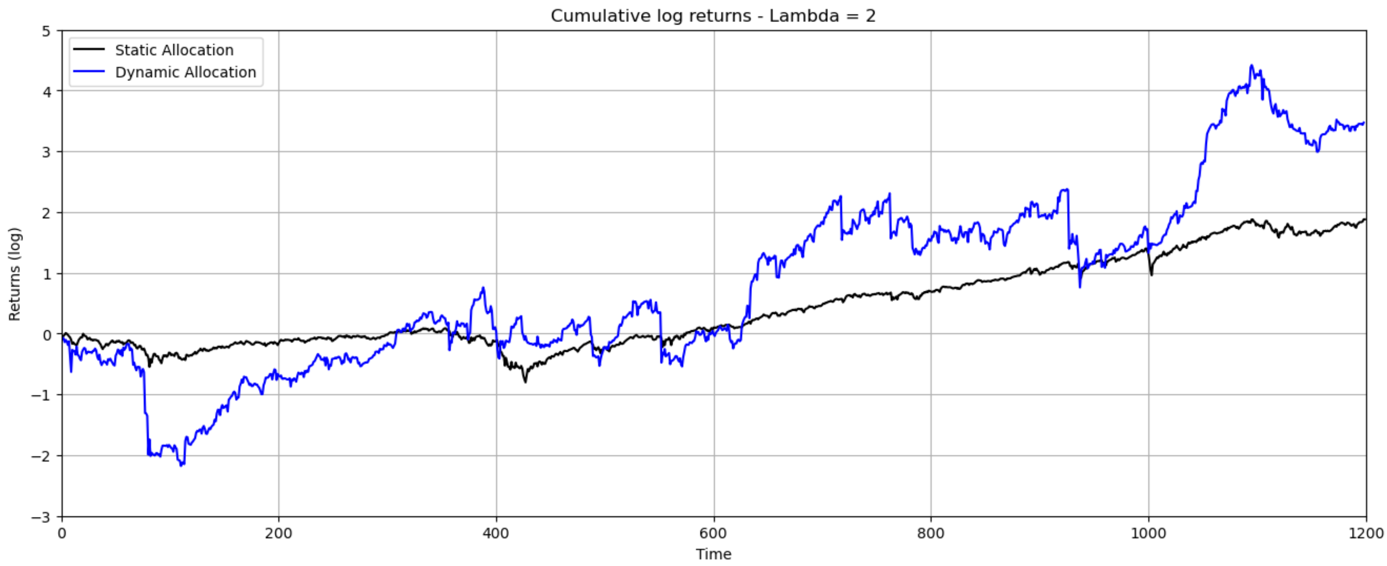


As mentioned previously, the weights for both assets are scalar multiples of one another for different values of the risk aversion parameter lambda.

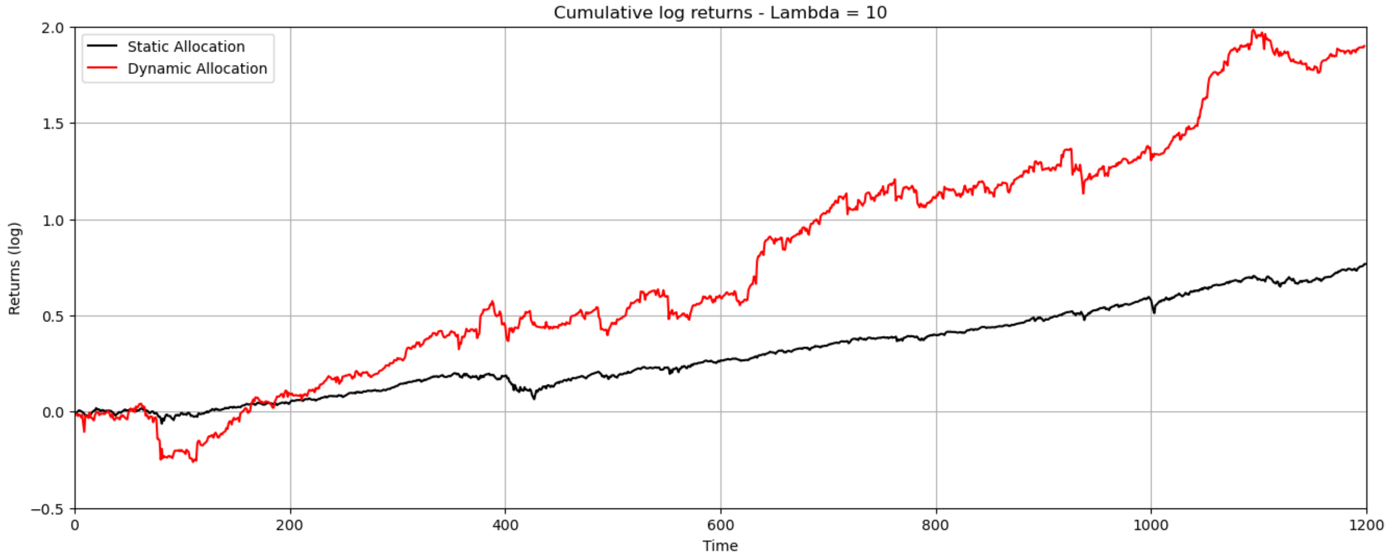
Q3.2

The cumulative returns for the $\lambda = 2$ strategy turned out to be explosive, therefore, we chose to use log returns for all strategies. Both dynamic strategies performed better in the long run than the static ones, and both were more volatile. The two $\lambda = 2$ strategies had periods of significant losses up until the middle of the sample, whereas the more conservative $\lambda = 10$ strategies saw comparatively tame losses, only at the beginning of the sample.

The $\lambda = 2$ static and dynamic strategies achieved a final cumulative log return of 1.90 and 3.47 respectively:



The $\lambda = 10$ static and dynamic strategies achieved a final cumulative log return of 0.77 and 1.90 respectively (note the difference in scale):



These graphs demonstrate the fact that the pairs of strategies for different risk aversions are almost exact scaled multiples of one another, and move somewhat conjointly, but at different scales. Various performance metrics are summarized below:

	Cumulative Return	**Volatility**	**Sharpe Ratio**
Highest	**Dynamic ($\boldsymbol{\lambda = 2}$)** **3.4723**	Dynamic ($\lambda = 2$) 0.0920	**Dynamic ($\boldsymbol{\lambda = 10}$)** **0.1127**
Lowest	Static ($\lambda = 10$) 0.7668	**Static ($\boldsymbol{\lambda = 10}$)** **0.0058**	Static ($\lambda = 2$) 0.0684

The best strategy in terms of cumulative performance was the dynamic $\lambda = 2$ strategy, although it also was the riskiest and thus only appropriate for risk seeking investors. For investors wishing to minimize risk, the static $\lambda = 10$ strategy offered the lowest volatility. The dynamic strategy with $\lambda = 10$ offered the best risk adjusted returns, with a good balance of risk and return. Thus, the factor determining which strategy is the 'best', is how risk averse we are.

Q3.3

The transaction cost parameter f that equates the total cumulative returns (TCR) of strategies 1 (dynamic) and 2 (static) must result in transaction costs TC_t that satisfy:

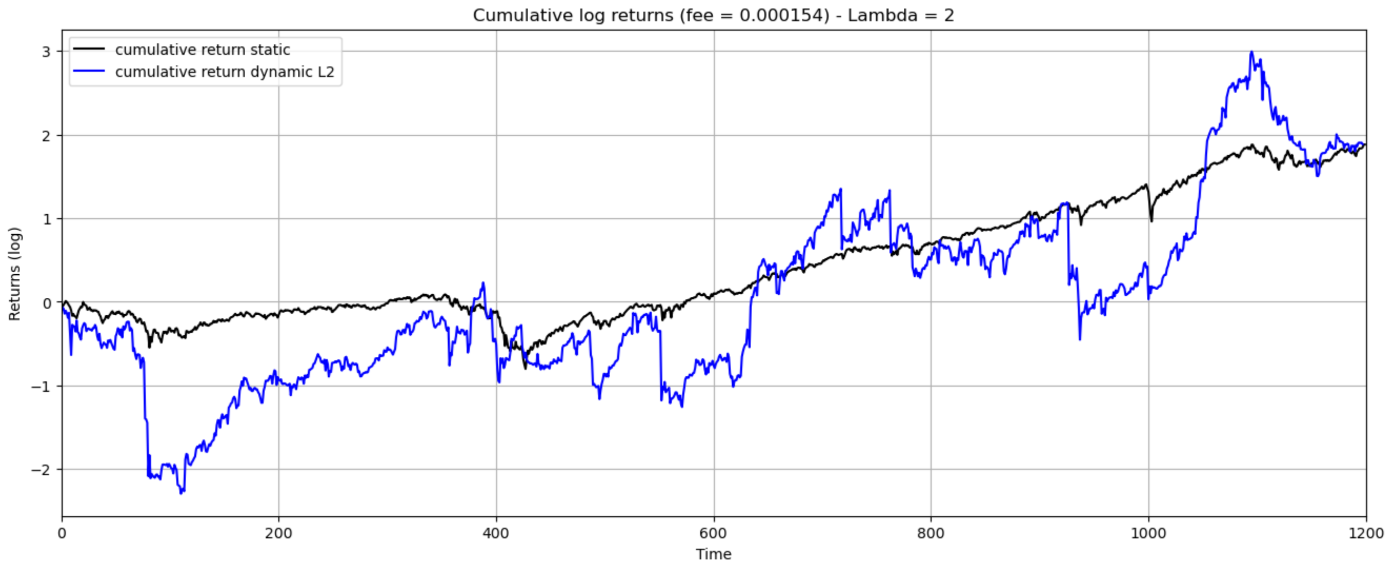
$$TCR^{(1)} = TCR^{(2)} \quad \text{or} \quad TCR^{(1)}_{\log} = TCR^{(2)}_{\log}$$

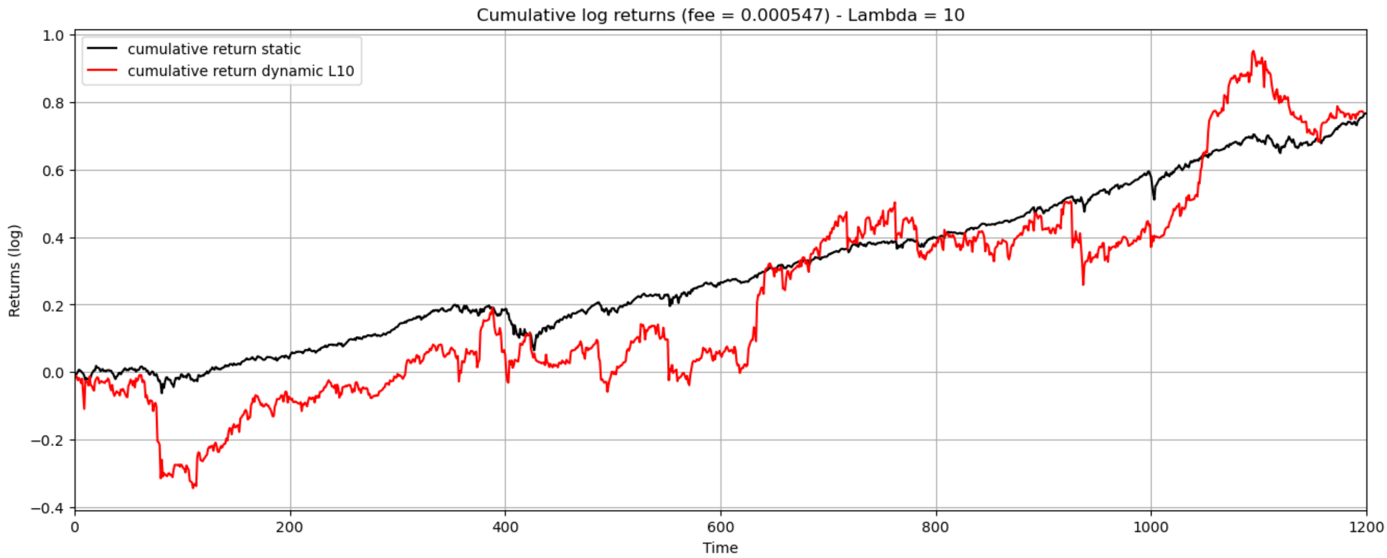
$$\left(1 + R_t^{(1)} - TC_t\right) \times \ldots \times \left(1 + R_T^{(1)} - TC_T\right) = TCR^{(2)} \quad \text{or} \quad \log\left(1 + R_t^{(1)} - TC_t\right) + \cdots + \log\left(1 + R_T^{(1)} - TC_T\right) = TCR^{(2)}_{\log}$$

We opted to use numerical optimization to determine the value of f that would satisfy the above equality. The optimal values of f that equate the returns for the two strategies to the static strategies were:

	$\boldsymbol{\lambda = 2}$	$\boldsymbol{\lambda = 10}$
$\boldsymbol{f}$	0.000154	0.000547

Interestingly, to match the performance of the static allocation, the $\lambda = 2$ strategy could absorb a fee that was lower than the $\lambda = 10$ strategy, despite the fact that it had a higher overall total cumulative return. This is explained by that the $\lambda = 2$ strategy had much higher weights in absolute terms, and hence much higher turnover.





[Part 4 below]

PART 4 – Computing the VaR of a Portfolio

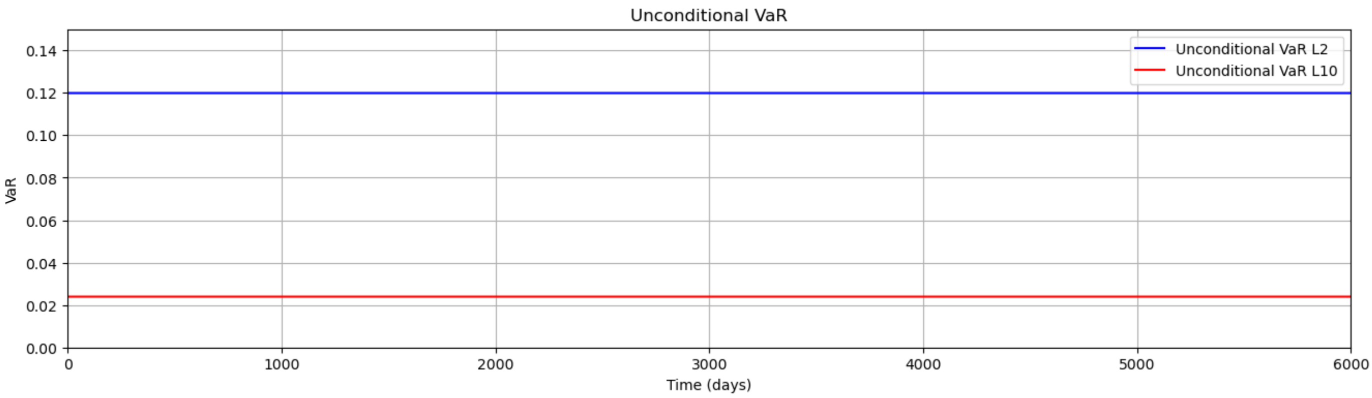
The general idea for this section is to model the distribution of extreme losses and the temporal evolution of the conditional 99% VaR by using the *extrema approach.* To do so, we must first determine the extreme distribution of the standardized quarterly maxima, and then use the quantiles of their distributions to infer the distributions of daily losses. We first define the daily loss as the negative returns $\left(L_{p,t+1} = -R_{p,t+1}\right)$, since we will be looking at the maxima. We use the static portfolio weights and the daily dynamic portfolio weights (weekly weights repeated 5 times) as demanded.

Q4.1

The unconditional Value-at-Risk (VaR) is calculated simply based on the sample mean ($\bar{L}_p$) and variance (σ_p^2), as well as the quantile corresponding to the probability $\theta = 99\%$ of a standard normal distribution: $z_{99\%} = 2.326$. The unconditional VaR ($VaR_p^{(Uncond)}$) is therefore computed as $\bar{L}_p + \sigma_p^2 \times z_{99\%}$ for both λs. The summary statistics are presented below in Table 4.4.1 along with its (lack of) evolution.

Table 4.1.1 : Unconditional VaR

	Unconditional Mean ($\bar{L}_p$)	Unconditional variance (σ_p^2)	Unconditional VaR ($VaR_p^{(Uncond)}$)
$\lambda = 2$	-0.001517	0.002724	11.991%
$\lambda = 10$	-0.000340	0.000109	2.3943%



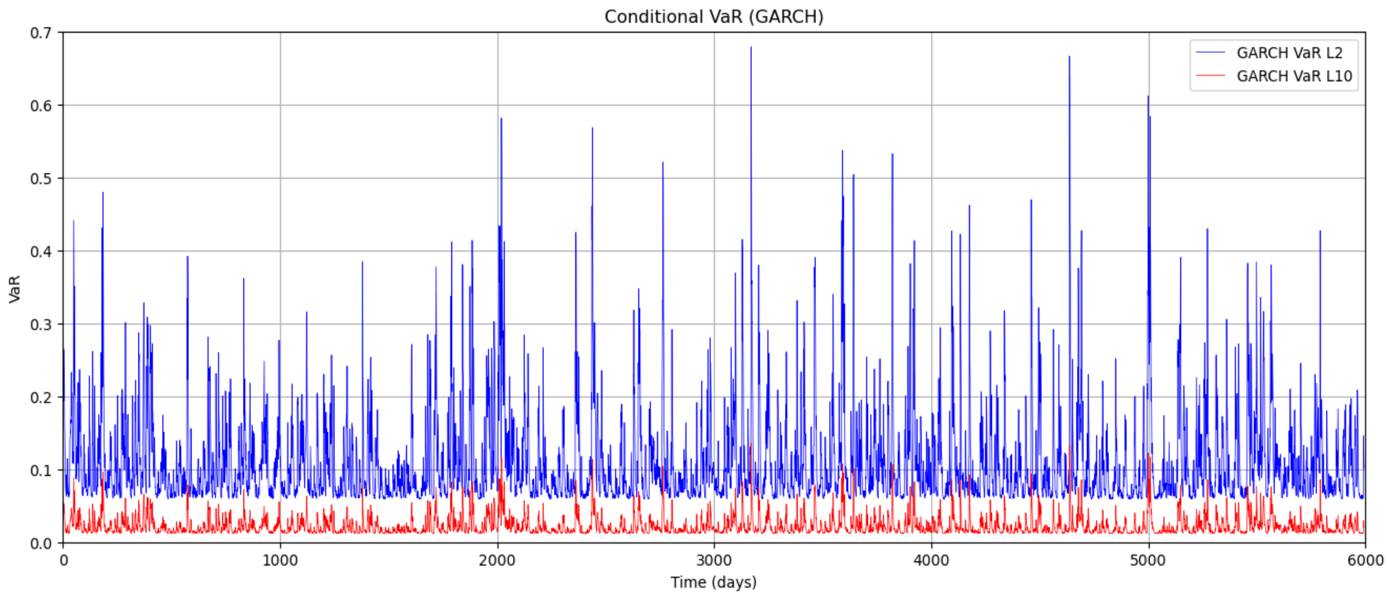
Q4.2

This section essentially repeats the GARCH estimation procedure as in Q2.3, except this time, we consider the volatility dynamics of the portfolio loss of the dynamic strategies with different lambdas, rather than the volatility dynamics of a single asset. The parameter estimates are presented below in Table 4.1.2. Comparing to Table 2.3.1, looking at portfolios instead of individual assets, the α is higher this time, suggesting that conditional variance of the loss is more sensitive to the shocks in the previous period. A lower β indicates that the volatility is less persistent this time. Moreover, $\alpha + \beta$ are almost identical between different λs and as compared to the GARCH estimation in Part 2, suggesting that the overall volatility persistence of assets is similar to the overall volatility persistence of portfolios of assets.

Table 4.2.1: GARCH Estimates for $L_{p,t+1}$

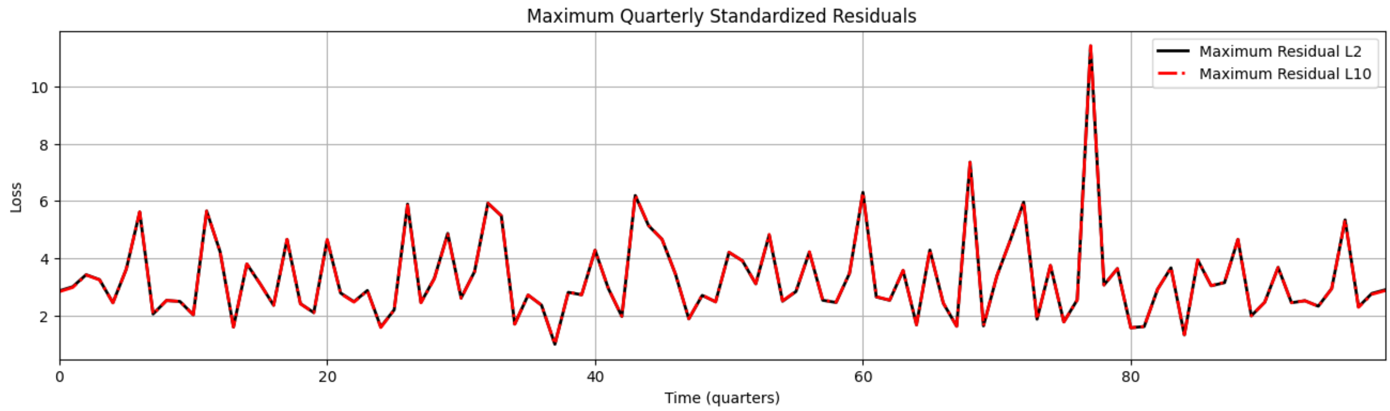
	$\boldsymbol{\omega}$	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\alpha + \beta}$
$\boldsymbol{\lambda = 2}$	0.000329	0.44282	0.519014	0.961830
$\boldsymbol{\lambda = 10}$	0.000013	0.44273	0.519324	0.962058

To compute conditional value at risk $VaR_{p,t+1}^{(GARCH)}$ for both λs, we need to estimate 3 new elements. First, we estimate the parameters for an AR(1) model for the daily losses. Second, we compute the expected losses by feeding the newly estimated AR(1) parameters to the formula $\mu_{p,t+1} = \alpha + \rho L_{p,t}$ where $L_{p,t}$ is the realized daily loss. Third, the variance forecasts are estimated using the new GARCH(1,1) parameters. The $VaR_{p,t+1}^{(GARCH)}$ is thus computed using the loss and variance forecasts $\left(\mu_{p,t+1} + \sigma_{p,t+1}^2 \times z_{99\%}\right)$ and assuming conditional normality for each period. The graph below shows the evolution of the expected losses, in the worst 1% of scenarios at any given point in time, represented by the value of $VaR_{p,t+1}^{(GARCH)}$ on the y-axis.



Q4.3

The modeling of the VaR using the generalized extreme value distribution (GEV) requires a strong assumption of X_t being a sequence of i.i.d. random variables, which is typically not the case for financial returns (or losses). Therefore, we focus on the standardized residuals instead, because the standardized residuals using the loss and conditional variance forecasts from AR(1)-GARCH(1,1) models $\hat{z}_{p,t} = \left(L_{p,t} - \hat{\mu}_{p,t}\right)/\hat{\sigma}_{p,t}$ should satisfy the i.i.d. assumption. We computed and segregated the daily standardized residuals into 100 quarterly subsamples (60 days each). For each sub-sample, we identified one maximum standardized residual and stored these 100 quarterly standardized residuals for later computations. Since the two strategies are almost exact scalar multiples, it is not surprising to see that their standardized residuals follow each other exactly, as depicted in the figure below:



Next, we procced to the estimation of the GEV parameters. Given the provided definition of generalized extreme value distribution (GEV) takes the following form:

$$H_{\xi,\varpi,\psi}(m_\tau) = H_\xi\left(\frac{m_\tau - \varpi}{\psi}\right) = \exp\left(-\left(1 + \xi\,\frac{m_\tau - \varpi}{\psi}\right)^{-1/\xi}\right)$$

We must first decide which probability density function (pdf) the GEV transforms to, then proceed to estimating the parameters of the GEV using maximum likelihood (ML) function. We expect the pdf of the GEV to follow the Fréchet or Weibull distribution where $\xi \neq 0$, which takes the form below:

$$h_{\xi,\varpi,\psi}(m_\tau) = \frac{1}{\psi}\left(1 + \xi\,\frac{m - \varpi}{\psi}\right)^{(-1/\xi - 1)} \exp\left(-\left(1 + \xi\,\frac{m - \varpi}{\psi}\right)^{(-1/\xi)}\right)$$

At last, maximizing the likelihood function:

$$L_T(\xi, \varpi, \psi) = \prod_{\tau=1}^{T/60} h_{\xi,\varpi,\psi}(m_\tau)$$

We took special care to ensure that an appropriate pdf of GEV is used throughout the optimization process. This was achieved by applying *'if'* conditions that would check whether the absolute value of the tail parameter ξ was below 0.05. If $-0.5 < \xi < 0.5$, then we would rerun the optimization assuming that the quarterly maximum standardized residuals follow the Gumbel distribution instead, using the corresponding pdf of GEV for the ML function.

Table 4.3.1: GEV parameter estimates for standardized residuals ($\hat{z}_{p,t}$)			
	ξ	ϖ	ψ
$\lambda = 2$	0.1181	2.6136	0.9819
$\lambda = 10$	0.1196	2.6119	0.9803

The Table 4.3.1 summaries the GEV parameter estimates for different λs. The shape parameters (ξ) are positive, which is to be expected as the financial series typically exhibits fat-tailed characteristics, implying a bigger likelihood of extreme values compared to a normal distribution. The location parameter (ϖ) indicates the central point for the maximum distribution of the standardized residuals, for both λs, they are very close to each other, indicating a stable (well estimated) central point. The scale parameters (ψ) indicate that the variability of the quarterly maxima of the standardized residuals are consistent across different λs, suggesting that the dispersion of the maxima is stable as well.

Q4.4

Now, we have all the parameters to compute the 99% quantile for the distribution of the maxima (m_τ) and the standardized residuals ($\hat{z}_{p,t+1}$). The quantile of the maximum distribution is computed using the following formula:

$$\hat{q}_\theta = \hat{\varpi} + \frac{\hat{\psi}}{\hat{\xi}}\left[(-\log(\theta))^{-\hat{\xi}} - 1\right] \text{ where } \theta = 99\%$$

Plugging in the parameter estimates from the Table 4.3.1, we obtain the 99% **quantile for extrema distribution** is **8.614%** and **8.625%** for $\lambda = 2$ and $\lambda = 10$ respectively. The interpretation is that with probability 1%, the maximum exceeds 8.614% (8.625%) for the $\lambda = 2$ ($\lambda = 10$) portfolio.

For the quantile of the loss distribution, we use a similar formula:

$$\hat{q}_\theta = \hat{\varpi} + \frac{\hat{\psi}}{\hat{\xi}}\left[(-N \times \log(\theta^*))^{-\hat{\xi}} - 1\right] \text{ where } \theta^* = 99\%$$

Again, plugging in the estimated parameters, we obtain the 99% **quantile for loss distribution** is **2.609%** and **2.607%** for $\lambda = 2$ and $\lambda = 10$ respectively. The same interpretation applies, with probability 1%, the loss exceeds 2.609% (2.607%) for portfolio with $\lambda = 2$ ($\lambda = 10$).

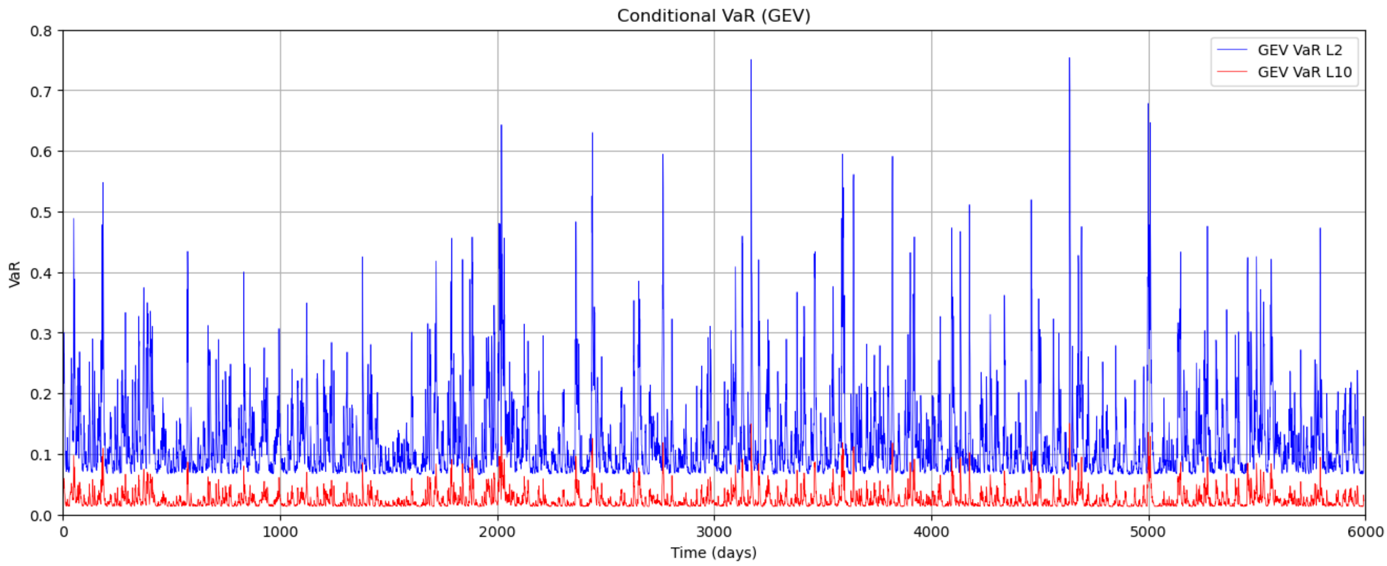
Q4.5

To compute the evolution of the 99% VaR of the loss distribution, the steps are essentially the same as for the $VaR_{p,t+1}^{(GARCH)}$, except we input the quantiles computed from the loss distribution for the two approaches.

$$VaR_{p,t+1}^{(GEV)} = \mu_{p,t+1} + \sigma_{p,t+1}^2 \times \hat{q}_{99\%}$$

The $VaR^{(GEV)}$ is generated for the entire sample, the plot below shows the evolution of $VaR_{p,t+1}^{(GEV)}$ for both strategies, with higher risk aversion (represented by the red line), the loss in the worst 1% case is much reduced when compared to the risk-seeking portfolio (represented by the blue line).

[Figure below]

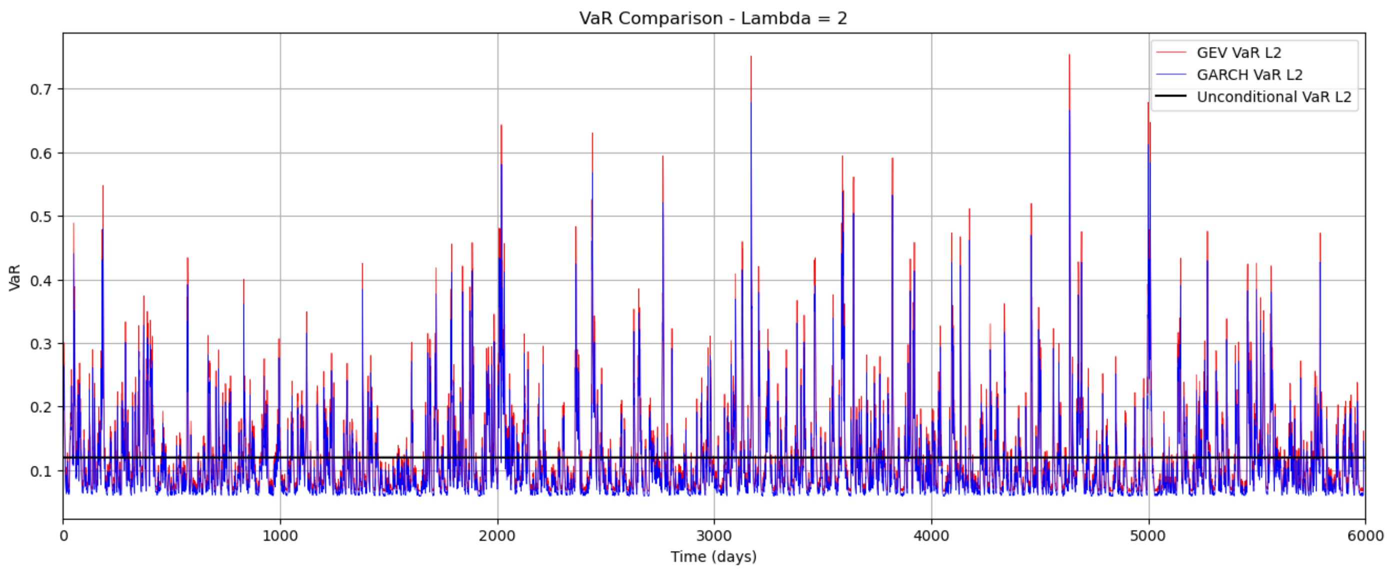


Q4.6

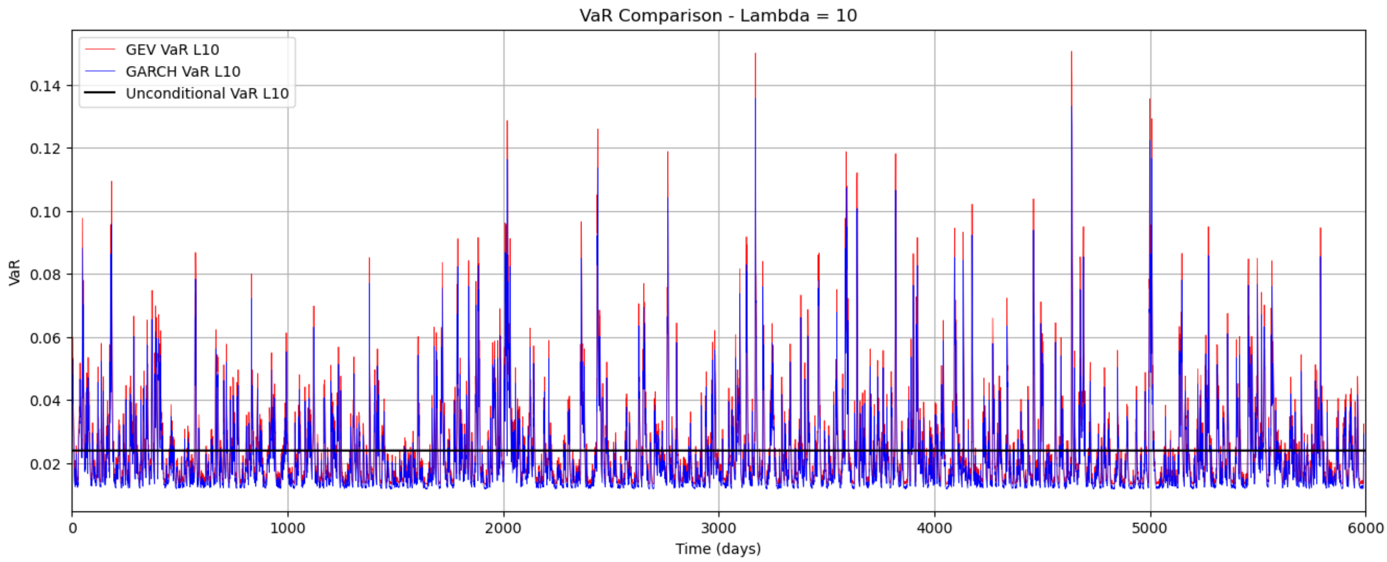
Comparing the 3 approaches, we can see that $VaR_{p,t+1}^{(GARCH)}$ and $VaR_{p,t+1}^{(GEV)}$ follow each other closely. The key difference between them is observable at times when the value at risk spikes. In those instances, $VaR_{p,t+1}^{(GEV)}$ suggest higher values. The fact that $VaR_{p,t+1}^{(GEV)}$ exceeds $VaR_{p,t+1}^{(GARCH)}$ suggests that in times of extreme volatility, the conditional VaR provided by GARCH systematically underestimates the strategy risk, since it assumes conditional normality. This result highlights the usefulness of using GEV quantiles for VaR modeling, because they provide a more faithful fat-tailed and volatility clustered representation of the distribution of extreme values that is often observed when looking at real financial series. The unconditional $VaR_{p}^{(Uncond)}$, being the least sophisticated of the three measures, offers limited precision in modeling the value at risk. It both underestimates the magnitude of expected losses in the periods of high volatility and overestimates them in the periods of low volatility.

For investors, having the 'correct' measure of risk is important because underestimating risk can lead to overexposure and significant losses, whereas overestimating the risk can lead to excessive caution, allocation to the risk-free asset, and unrealized returns.

These dynamics are presented in the following graph for the $\lambda = 2$ strategy:



For the $\lambda = 10$ strategy, the VaR also appears to be almost an exact scalar multiple of the $\lambda = 2$ strategy. This can be seen by noticing that the highest VaR value (0.15) for the $\lambda = 10$ strategy, is equal to the highest VaR value (0.75) of the $\lambda = 2$ strategy scaled by the ratio of lambdas: $0.75 * (2/10) = 0.15$. With that in mind, the VaR dynamics of the $\lambda = 10$ are presented below:


VaR Comparison - Lambda = 10
GEV VaR L10
GARCH VaR L10
Unconditional VaR L10
0.14
0.12
0.10
0.08
0.06
0.04
0.02
VaR
0
1000
2000
3000
4000
5000
6000
Time (days)

[The END]