POLITECNICO
MILANO 1863

Financial Engineering

Assignment 4 Cross-Correction
Group 16

2023-2024

Alice Vailati - CP: 10683600 - MAT: 222944
Andrea Tarditi - CP: 10728388 - MAT: 251722
Jacopo Stringara - CP: 10687726 - MAT: 222456
Nicolò Toia - CP: 10628899 - MAT: 247208

Contents

1. Introduction

After reviewing our code we set our results as a benchmark against which to evaluate other groups' work. In the following work we give our evaluation of these groups and suggest a final mark. We also explain, in sufficient detail, the procedures and methodologies we adopted in order to answer the questions ourselves and consequently evaluate our peers.

2. Variance-Covariance Method for VaR and ES with t-student distribution

In this first exercise our aim is to compute the VaR and ES at the $\alpha = 99\%$ level of an equally weighted portfolio made up of the following:

- Adidas
- Allianz
- Munich Re
- L'Oréal

by employing a variance-covariance method on a t-distribution with $\nu = 4$ degrees of freedom. The notional of our portfolio is 15 million euros, the data we used were those available between the valuation date, 20th of February 2020 and 5 years prior.
First of all we can linearly approximate the losses using using the log-returns and their respective weights. Hence:

$$L(\underline{X}_t) = -\sum_i \omega_i X_{i,t}$$

Then we compute its mean μ and variance σ^2 using the mean vector and variance-covariance matrix of the log returns and the weights.
Then the VaR can thus be obtained as:

$$VaR_\alpha(L) = \mu + \sigma t_\nu^{-1}(\alpha)$$

while the ES can be computed as the following:

$$ES_\alpha(L) = \mu + \sigma ES_\alpha^{std} \quad \text{with} \quad ES_\alpha^{std} = \frac{\nu + (t_\nu^{-1}(\alpha))^2}{\nu - 1} \cdot \frac{\phi_\nu(t_\nu^{-1}(\alpha))}{1-\alpha}$$

These computations lead to:

VaR	ES
563223.32	787977.23

Table 1: VaR and ES in euros at 99% confidence level

We recall that VaR at level $\alpha = 0.99$ is a risk measure indicating a potential loss that is exceeded with probability 1%. While the ES is the average of values beyond the VaR, indeed ES is always greater than the VaR.
On a notional of 15 million euros, we stand to exceed a loss of 560k euros (about 3.75%) in only 1% of scenarios each day.
Since there are about 256 business days in a calendar year, we may expect to incur in a loss greater than the VaR about 2.5 times in this period. This means we may expect a relative loss of more than 1.4 million euros in the 1% scenario over a year's time. Although this value is relatively high, this can be understood by the composition of our portfolio which has a 100% exposure to equity.

2.1. Group 12

Group 12 did a great job in point zero. They delivered very readable code, pointing out key elements in the report. Moreover, the results provided are coherent with what we expected and found ourselves. Comments in the code are up to the point and not overbearing. The interpretation and explanation of the results and methods used to obtain them are more than satisfactory.
We awarded them full marks in this exercise.

2.2. Group 13

Group 13 chose to use simple returns of the stocks, i.e. $\frac{S_{t_i}}{S_{t_{i-1}}} - 1$, rather than the log returns. This fundamental error then pollutes all their work in points 0 through 1.
From a quantitative prospective adopting simple returns instead of log-returns is a critical mistake, showing a flawed understanding of basic concepts. The main advantage of log returns is that we can easily aggregate them across time, unlike simple returns (this is known as the additivity property). Log returns are symmetric around 0 and their values can range from minus infinity to plus infinity whereas simple returns' downside is limited to -100%.
A negative movement of -25% (e.g. from 100 USD to 75 USD) does not reverse the losses by going +25%. (75 USD to 93.75 USD). In this particular case the choice of simple returns can be partially "justified", thanks to the asymptotic behaviour of the returns. Indeed, since we are working with daily returns with relatively small values and for time intervals extended enough, the returns and log-returns should coincide and the additivity property should hold as well. However, it is crucial to say that this is not true in general, for returns computed annually, quarterly, monthly or even weekly the returns and log-returns do not coincide asymptotically.

Aside from this choice their implementation of the t-student parametric approach for the VaR seems correct. For the ES, on the other hand, the standard ES is multiplied for the rescaling factor Δ rather than for $\sqrt{\Delta}$, but thanks to the fact that in our case we had $\Delta = 1$ this does not impact the result.
No reason is given on why the results are actually coherent with the notional of 15 million euros.
In our honest opinion, the material provided for this point is unsatisfactory even though their results are relatively close to the correct ones thanks to the asymptotic property of simple returns.

2.3. Group 14

Group 14 did a good job with the coding part, commenting in a proper way both the functions and the main file, obtaining the same results as us and Group 12. However, even though their paper properly explains the methodology employed it completely lacks elucidation regarding the obtained results.
We deemed Group 14's work on this point satisfactory from a practical point of view.

3. Historical simulations, bootstrap and PCA approaches to VaR and ES

In this part of the assignment we employ a variety of techniques to compute the VaR and ES of various linear portfolios. All computations are done on data available between the 20th of March 2019 and 5 years prior. For this part all VaR and ES are computed with a confidence level of $\alpha = 0.95$. Furthermore, throughout this case study we operated under the 'frozen portfolio' assumption, that is, that the portfolio weights are constant across the time horizon taken into consideration.

3.1. Historical Simulation & Bootstrap

In this part of the case study we take a look at two non-parametric approaches to compute the VaR and ES, focusing on the Historical Method and the statistical bootstrapping technique.
In particular we will apply them to the following portfolio:

- Total (25K shares)
- AXA (20K Shares)
- Sanofi (20K Shares)
- Volkswagen (10K Shares)

for a total portfolio value of 4731416.99 euros.
In order to apply the historical method we compute our losses using a linear approximation just like for point zero, and subsequently order them in decreasing order:

$$L^{(n,n)} \leq \cdots \leq L^{(1,n)}$$

Then the VaR and ES can be computed as follows:

$$VaR_\alpha = L^{(\lfloor n(1-\alpha)\rfloor,n)}$$
$$ES_\alpha = mean\{L^{(i,n)} : i = \lfloor n(1-\alpha)\rfloor, \ldots, 1\}$$

To then apply the statistical bootstrap method we extracted 200 rows of our log returns with replacement and applied the same technique as above on the bootstrapped dataset. It is good practice to fix the seed of the

random generator in order to obtain consistent results across different runs of our code for the bootstrapping technique.
These two techniques yield the following values:

	VaR	ES
Historical Method	96039.47	143630.83
Bootstrap	100759.30	174036.41

Table 2: VaR and ES expressed in euros

In order to check our results we apply the well known thumb rule of the plausibility check:

$$l_i = VaR_{X_i}(1-\alpha) \quad u_i = VaR_{X_i}(\alpha)$$
$$sVaR_{X_i} = sens_i \frac{|l_i| + |u_i|}{2}$$
$$VaR^{ptf} = \sqrt{sVaR \cdot C \cdot sVaR}$$

where C is the correlation matrix of the risk factors. The plausibility on this portfolio yields a value of 92035.63 euros. Relying on the thumb rule, which states that the model VaR can be deemed correct when it is of the same order of magnitude as the approximated VaR, we can deem our results trustworthy.

3.2. Weighted Historical Simulation

We used this alternative non-parametric approach to estimate the risk associated to an equally weighted portfolio made up of Adidas, Airbus, BBVA, BMW and Deutsche Telekom.
In particular, for the weights we used a decay factor $\lambda = 0.95$. To apply the weighted historical technique we computed the losses of the portfolio using a linear approximation and ordered them in descending order just like we did for the simple historical approach above. The main difference here is how we estimated the quantile for the VaR, in particular, given the following weights:

$$w_s = C\lambda^{(t-s)} \qquad \text{with } C = \frac{1-\lambda}{1-\lambda^n}$$

We have that the VaR and ES are computed as follows:

$$VaR_\alpha = L^{(i^*,n)} \text{ where } i^* \text{ is the largest value such that } \sum_{i=1}^{i^*} w_i \leq 1-\alpha$$
$$ES_\alpha = \frac{\sum_{i=1}^{i^*} w_i \cdot L^{(i,n)}}{\sum_{i=1}^{i^*} w_i}$$

Applying this technique yields the following results:

VaR	ES	Plausibility Check
1.593678%	2.154424%	1.921928%

Table 3: VaR and ES as a percentage of portfolio value (here equal to 1)

We can easily see that the Plausibility check has the same order of magnitude as the VaR.
Let us remark that weighing our observation not solely on their magnitude but also on their recentness, let's us obtain a more accurate result of the actual risk the portfolio possesses. Indeed, in general the WHS method is less sensitive to far-in-the-past black swan events which may otherwise greatly influence our result but are here mitigated through the weighting process and the decay factor λ.

3.3. PCA approach

We now pass to evaluating the VaR via a parametric approach: the Principal Component Analysis method.
This method, unlike the previous two, has a further assumption of normality. In particular we must assume that the factors obtained via the PCA are normal.
Here we are trying to measure the VaR of a portfolio made up of the first 18 companies in our dataset. Thus it becomes quite clear that reducing the number of factors used in our estimation with a dimensionality reduction

technique such as PCA could improve our estimation of the VaR. In particular, the PCA approach lets us recombine the previous features into new, independent factors and moreover, the factors are ranked by how much of the original variance they explain. Thus, restricting ourselves to a smaller number of factors has the added benefit of separating the real data from the underlying noise in the observations.
To apply the PCA approach we estimated the eigenvalues and eigenvectors of the yearly variance-covariance matrix (since the variances are of the same order this is fair) using the built-in numpy function. Then we translated the original weights and mean into the new basis:

$$\underline{\hat{\mu}} = \Gamma^T \underline{\mu} \quad \underline{\hat{w}} = \Gamma^T \underline{w}$$

Then we can easily apply the PCA by computing the reduced mean and variance for any $k \in [1, N]$ as follows:

$$\sigma_{red}^2 = \sum_{i=1}^{k} \hat{w}_i^2 \lambda_i \quad \mu_{red} = \sum_{i=1}^{k} \hat{w}_i \hat{\mu}_i$$

Then the VaR and ES can be computed by taking a Gaussian parametric approach on these reduced mean and variance:

$$VaR_\alpha = \mu_{red} + \sigma_{red} \cdot \mathcal{N}^{-1}(\alpha)$$
$$ES_\alpha = \mu_{red} + \sigma_{red} \cdot \frac{\phi(\mathcal{N}^{-1}(\alpha))}{1-\alpha}$$

This yields the following:

n	VaR	ES	Explained variance (%)
1	5.453693%	6.881475%	52.66
2	5.431908%	6.861275%	59.75
3	5.434004%	6.863389%	64.28
4	5.441210%	6.870956%	68.72
5	5.439734%	6.869500%	72.77

Table 4: VaR and ES computed as percentages of portfolio value along with explained variance

As a final note we can remark that the new factors found via PCA are nothing else than linear combinations of the previous features. In our case this also has a nice financial interpretation: the new factors are nothing more than linear portfolios based on the stocks in the original portfolio. In particular, we may interpret the portfolio with the highest explained variance as the "market portfolio", indeed this portfolio, although modestly, is in a sense driving the variance of the original portfolio. Furthermore we may also give a financial interpretation to the fact that the VaR slightly decreases when passing from considering $n = 1$ to $n = 2$ factors. This means that the portfolio with the second highest eigenvalue is a portfolio that works against the market, we may call it the "antithetic" portfolio.
On a final note we encourage to use $n = 5$, indeed with PCA one usually tries to achieve at least 70% to 80% of the original variance. The fact that a relatively high number of factors must be used to achieve this is probably due to the fact that all the companies we are taking into consideration are part of the Eurostoxx50 Index and hence are highly correlated and tend to move with the market (or rather, they move the market).
In conclusion, referring to the G.Pasini paper: *"It is worth to mention that such type of analysis do not cover the investor against hidden financial risks, as those related to deteriorated financial positions of the company's stocks involved in the PCA selection procedure, or against possible default contagion over financial networks."*

3.4. Group 12

In the introduction Group 12 clearly and thoroughly explain their assumption and methodology for the loss computation. In particular, we appreciated their level of attention to detail, for instance they explained with great clarity how and why they computed the portfolios' linearized losses, values and weights.
For what concerns the Statistical Bootstrap and Historical Simulation methods, they likewise explained themselves well and obtained results aligned with ours. Our only concern is the seed not being fixed in the main file prior to running the Bootstrap, leading to differing results across runs even though they themselves acknowledged this in their report.

Group 12 gave a very detailed introduction to the WHS method. They explain all formulas correctly, even including the rescaling factor $\sqrt{\Delta}$ for both VaR and ES. However their Expected Shortfall presents a minor mistake related to python indexing choosing an i^* that is off by one.
Their implementation of the PCA presented one mistake, they did not change the sign of the mean vector μ of the returns to match the one of the loss. Their results are thus a bit off the mark. Although they spent a lot of time detailing their methodology, they did not provide a solid reason on why a choice of $n = 3$ or $n = 5$ factors would be good.
Group 12 overall did a solid job in this point, well commenting and explaining what they did. Aside from minor mistakes they did a satisfactory job.

3.5. Group 13

Group 13 did not explain the assumptions under which they worked throughout this exercise. On the whole, exercise one contains multiple mistakes and points of failure.
Regarding their Historical Simulation, their first mistake was taking into consideration the wrong portfolio (they used the one from point zero) and weighing them by the number of shares rather than their value. As mentioned above in point zero, they also kept on using simple returns.
Their linearized loss is wrongly computed with no minus sign in front. Their report is completely lacking in showing the procedure followed.
Bootstrap was wrongly implemented, Group 13 did not use replacement. This highlights a lacking understanding of the underlying theoretical concept. Again, there are no comments on methodology and procedures.
Their plausibility check would give the correct result if fed the correct inputs. However, there is a misplaced minus sign which cancels out in the squaring procedure.
Group 13's implementation of the WHS and PCA procedure seems correct although they again performed it on the simple returns thus giving an incorrect output. Their report, again, lacks any meaningful explanation or comments.
Overall we, as a team, were unable to debug their code. Its convoluted structure hindered our best efforts to discern what Group 13 correctly implemented and what not. As a consequence of these multiple points of failure, their provided results are overall simply wrong.
Group 13's work is marred by their initial choice of simple returns. Although, some of their function are actually implemented in the correct manner, their work remains overall inadequate.

3.6. Group 14

Group 14 encountered a few difficulties with the Historical Simulation section. Specifically, while utilizing their *relevant_returns* function, they generated a matrix with columns ordered alphabetically. However, they subsequently multiplied these returns by a vector of weights still in its original order. Consequently, this error led to the erroneous consideration of 25k shares of AXA and 20k shares of Total Energies, instead of vice versa. If the correct order is preserved, Group 14's work gives results aligned with ours and Group 12's. Although their results use the correct procedure for $\Delta = 1$, they do not rescale the VaR and ES for the time horizon at all in their functions. Hence their function works in this particular case but non in general. Their bootstrap is correctly implemented, we encourage them to use built-in functions (such as pandas sample data-frame method) rather than extracting indexes directly. For what concerns their report, they were concise to a fault. Indeed, they skimmed the surface of the Historical Method and Bootstrap procedure without delving into details.
The parts concerning the Weighted Historical Simulation is implemented correctly except for the missing rescaling factor. The results align with the correct ones, although their report is rather meager.
Their PCA is be implemented in the correct manner and thus they obtain the correct results. It is worth noting that they opted not to evaluate the percentage of Explained Variance in the PCA, which we believe is one of the most significant outcomes deducible from the application of this computational method. They also chose not to elaborate further on which choice of the number factors would be best.
On a final note their plausibility check was correctly implemented and well explained and to the point.
Group 14, aside their small mistake in the HS part, performed rather well in the practical part of the assignment but their report lacks in interpretation and elaboration on the results.

4. Full Monte Carlo and Delta-Normal VaR

In this case study we pass from evaluating the VaR for simple linear portfolios to non-linear portfolios, i.e. portfolios that contain non-linear payoff derivatives, such as calls.
In this specific case we were asked to determine the 10 day VaR at the 95% confidence level on the 16th of January 2017 of a portfolio consisting of stocks of BMW for 1,186,680 Euro and shorting the same number of

call options on the stock with expiry on the 18th of April 2017, with strike 25 Euros.
For the call we have a volatility of 15.4%, the stock also has a dividend yield of 3.1% and there is a fixed interest rate of 0.5%.
In order to price a non-linear derivatives portfolio we used two different techniques: Full Monte-Carlo and a Delta normal approach. In both cases we simulated the movement of the underlying with a Weighted Historical Approach, in other words, to make our simulations we extracted N_{sim} observations with replacement from the last 2 years having a probability to extract each observation equal to $w_s = C\lambda^{(t-s)}$ with $C = \frac{1-\lambda}{1-\lambda^n}$ and $\lambda = 0.95$. Then we simulated the price at time $t + \Delta$ as the exponential of the returns extracted above re-scaled by Δ:

$$S_{t+\Delta} = S_t \cdot e^{x_t \cdot \Delta}$$

This passage is shared by both methods, they differ in how we then compute or approximate the loss. For both cases we supposed the Black & Scholes dynamics for the underlying. Then we did the following:

- For the Full Monte Carlo approach we computed the price of the Black and Scholes formula at time t_0, which is the limit in mean of running a full Monte Carlo simulation, and then we also computed the call price for each simulation at time $t_0 + \Delta$, i.e. we computed it with the new price $S_{t_0+\Delta}$ and a time to maturity decreased by Δ. Let us remark that in this case, by assuming Black & Scholes dynamics for the underlying we can skip running a true full Monte Carlo since we already have the closed formula to which the mean of the simulations will converge.
- For the Delta normal technique we chose to approximate the change in the price of the Call by using its Greek with respect to the underlying price, the Δ. In order to do this we computed the Black & Scholes delta for the call and multiplied it by the change in price of the underlying.

In both cases we computed the loss of the asset simply as its change in value. Subsequently we computed the loss of the whole portfolio as the difference of the asset loss and the call loss, since we are shorting the option. Finally in both cases we ordered the losses from biggest to smallest along with their weights and computed the VaR just like we did in the WHS approach.
These two techniques yielded the following numerical results:

Full Monte Carlo	Delta Normal
4400.22	4155.89

Table 5: VaR for both techniques expressed in euros

4.0.1 Delta normal improvements

In general, since we are approximating the derivative of the call price, we can improve our results by using our estimates for other Greeks. For example we could adopt the well known Delta Gamma approach by also employing the Gamma in our estimation. In this particular case, however, this would not be very impactful. Indeed, we have a call option very close to its maturity and far in the money since $K = 25$ euros, hence we know from theory that its Black & Scholes Gamma would be close to 0 (we refer to 'Arbitrage theory in continuous time' by T.Bjork). Further improvements could be obtained by computing the Θ, the Greek with respect to time, or the ρ, the Greek with respect to the interest rate, thanks to the fact that these possess closed formulas and we could easily compute the change in time and interest rate.

4.0.2 On the cost of the Monte Carlo approach

Let us remark that in our case, we were able to bypass running a true, fully fledged nested Monte Carlo simulation thanks to the fact that we were able to employ a closed formula.
In case of more complex derivatives for which a closed formula does not exist, e.g. options whose payoff depends on the path of the underlying, one needs to run a nested Monte Carlo simulation in order to price the option at time $t_0 + \Delta$. In other words, in the case of exotic derivatives the number of simulations we need to run grows exponentially, especially if we have multiple factors of risk.

4.1. Group 12

Group 12 performed their full Monte Carlo simulation by drawing N_sim sets of 10 returns weighing, just like us, each observation by its WHS weight. Their approach differs from ours, in that we simply rescale the returns, while they extract 10 for each simulation and sum them up.

Their results are overall coherent and reasonable. Moreover they explained their choice and how they performed the simulation.
For the Delta Normal approach the results provided are a bit too far, in our opinion from their own results for the MC procedure. While they gave an explanation we did not find it convincing but, still, coherent with their modus operandi.
Their comments about the theoretical questions were quite comprehensive and solid.

4.2. Group 13

Group 13 chose, just like Group 12, a different approach than ours. Before drawing their sample, they reshaped the data into a rolling window form, then they drew samples from these rather than from the original observations. We found this novel approach very impressive. It has the potential to fully capture the clustering of returns across the period. Moreover, they correctly started to use log returns and fully motivated their modeling choices.
Just like Group 12, their Delta Normal result is quite far from the MC approach but, unlike the others, they clearly recognized the mismatch in the values and suggest potential ways to remedy this.
They argued their answers to the theoretical questions in a convincing manner.

4.3. Group 14

Group 14 wrongly performed a uniform drawing of the sample in the Monte Carlo approach, as they themselves indicated in their report the results are clearly wrong. No reason is given on why this could be.
What we wrote above also applies to their Delta normal approach but their practical result is even worse.
Their discussion on pros and cons of the approaches were quite meager, if not insufficient.

4.4. Disclaimer

In the Full Monte Carlo technique, we chose to accept more than one interpretation.
Our group decided to be more conservative overall and to simply re-scale the returns, so that negative observations weighed more significantly. Group 12, on the other hand, by summing different days arrive at a more "optimistic" result, indeed summing returns from different days tends to values around the mean and to mitigate the effect of the variance. While Group 13 decided to keep the time relation between days and obtains an estimate in between the two previous approaches.
We believe that all these methods are valid. Which is to be used, depends on the circumstances. For example, our approach for the VaR could be preferred when caution is needed, like in a regulatory framework, while Group 12 and 13's could be preferred from an operative point of view, because it could significantly lower the required capital (if approved).
Overall, we believe Group 13's approach takes the best of both worlds. It keeps the time relation between days and yields a very reasonable estimate.

5. Pricing in presence of counterparty risk: Cliquet option

Now, we pass to evaluate a final case study: the pricing of a Cliquet option in presence of counter party risk. We study the case of a cliquet option written on an underlying S_t and issued by our counterparty Intesa San Paolo.
Importantly we can think of the Cliquet option as a sum of ATM forward-starting option with each one having a payoff as below:

$$\left[L \cdot S_{t_i} - S_{t_{i-1}}\right]^+ \qquad \text{for } i = 1y, \ldots, 7y$$

Where L is known as the participation coefficient and is equal to 0.99.
In case of recovery we considered that we would be able to recover about $\pi = 40\%$ of the NPV at that time. Hence we can write the price at time t_0 as:

$$Price(t_0) = \mathbb{E}_0 \left\{ \sum_{i=1}^{N} D(0,t_i) \cdot \left[L \cdot S_{t_i} - S_{t_{i-1}}\right]^+ \cdot \mathbb{I}_{\tau > t_i} + \pi\, D(0,t_i)\, NPV(t_i) \cdot \left(\mathbb{I}_{\tau > t_{i-1}} - \mathbb{I}_{\tau > t_i}\right) \right\}$$

This lead to the following closed formula:

$$Price(t_0) = \sum_{i=1}^{N} \Bigg\{ B(t_0; t_{i-1}, t_i) \cdot S_0 \cdot P(t_0, t_i) \cdot C\left(L \frac{B(t_0, t_{i-1})}{B(t_0, t_i)}, 1, \sigma \right) +$$
$$+ \pi \ \cdot [P(t_0, t_{i-1}) - P(t_0, t_i)] \cdot S_0 \cdot \left[\sum_{j=i}^{N} C\left(L \frac{B(t_0; t_{j-1})}{B(t_0; t_j)}, 1, \sigma \right) \right] \Bigg\}$$

For a full derivation we refer the reader down below to the attached appendix.
To have a benchmark against which to test our formula we also performed a monte carlo simulation.
We simulated the time of default by drawing from a standard uniform and inverting the probability of default $P(0, t_i)$ (using the ones obtained from the CDS spreads of ISP) to find the time of default τ. We also simulated the underlying price by assuming Black & Scholes dynamics for it in 7 steps. In order to compute the discount factors at each required point in time we used the ones we bootstrapped previously.
For each simulation we then computed the required cash flows and computed the mean.
Applying both the closed formula approach and the Monte Carlo ($N_{sim} = 1000000$) we pervene to the following results:

Monte Carlo	Closed Formula
19436118.26	19438069.75

Table 6: Price for a cliquet option in euros and notional 30mln

The MC method has a 95% IC of [19409295.63541795 – 19462940.87582682] euros.
Thanks to the fact that with a relatively small number of simulations we were able to obtain a result that is remarkably close (it differs of only 2000 euros) to the MC one, we decided to trust our closed formula.
In general, the cliquet, like all other derivatives, presents no counterparty risk for the issuer, thanks to the fact that they have an always negative NPV.
Cliquet options are usually used to build more complex structured products and are usually traded over the counter. They are also bought by Insurers to mitigate risks. This leads to them having higher fees and transition & operational costs compared to traditional options or other investment products.
The price at which ISP would try to sell the cliquet option is the fair price plus the operational costs and fee, which would be relatively high in order to cover the exposure with simpler products (e.g. long EU call) and offset the cost of managing such an option. Moreover, the L coefficient grants space for ISP to gain in case the underlying goes above its previous value.

5.1. Group 12

Group 12's computed the value of the cliquet option via a closed formula and an hybrid approach, and their results seem quite coherent. However, the results of the two approaches differ in a non negligible way. This difference can probably be explained by a minor mistake in computing the recovery value. Indeed, while all other passages are clearly explained, how Group 12 computes the future cashflows is not clear.
Their comments, while formally correct, do not provide meaningfull insights.

5.2. Group 13

Group 13 only employed a closed formula. Their explanation of how to arrive at such a formula was too brief and rushed. They do not explain at all, how the recovery value is computed.
In our opinion, their pricing is off the mark.
Their explanation of at what price and when to sell a cliquet is quite generic and did not convince us.

5.3. Group 14

Group 14 in the preamble make a few strong (and wrongful) claims but correctly do not employ them in their derivations.
Their derivation is overall confusing and there is a clear lack of passages. Indeed, in the initial formula they forgot to discount the cash flows. Furthermore there is an explanation of how the recovery value should be taken into consideration but no formula is provided. Finally, to justify the huge discrepancy between their MC approach and their analytical result, they blame inaccuracies in the Black-76 model. However they successfully employed the AV technique to reduce the number of simulations for MC.
We found their comments about the Cliquet price quite unsafisfactory.

6. Final Evaluation and Comments

6.1. Group 12

Group 12 performed rather well.
Their results were mostly correct and their coding, while not optimized for python, performed its duty efficiently.
Where they really shined was in how they wrote their paper. We found it very well written and with many details. In some respects even superior to our own work.
Overall they showed a solid understanding of the theory concepts, they demonstrated proficiency in employing them and applying them to real world case studies.

6.2. Group 13

In the case of Group 13 close attention must be paid to their coding. In the preamble to their report, they clearly declared their intention to set out a python class framework to model portfolios and options in order to take full advantage of python OOP-oriented approach.
In their intentions this class framework would have helped to abstract away from the underlying data structures and help financial engineering code with more human friendly concepts.
Unfortunately, Group 13's coding falls short of their very ambitious goals and is far less readable and robust than they claim it to be.
We as a team found it quite hard to read and debug team 13's code and understand why and how certain architecture and design choices are taken is not entirely clear.
Let us make an example: the portfolio class implements various methods to compute the VaR and ES, indeed they implemented all the required functions. In their implementation a portfolio can be made up of both linear and non-linear assets but some of their method to compute VaR only take into consideration the linear part of the portfolio. While correct from a theorical point of view, it is not clear why the chose to aggregate these two types of portfolio into a macro class rather than developing two subclass, one for linear portfolios and one for non-linear portfolios.
Furthermore their work has many Python and OOP bad practices (indeed the Portfolio class does not fully implement data encapsulation). These range from using global variables for the number of days in a year, when it should really be a class attribute for the Portfolio class, to having public methods potentially adding and modifying attributes of the class at run-time.

Group 13's coding is commendable in spirit but short-sighted in practice. Their ideas show great potential to develop into a fully fledged risk management library. However, before this can be achieved, a closer look must be paid to their works foundational assumptions and design choices in order to have a strong foundation on top of which to build their framework.
Other than this, their code is actually well commented and relatively readable and their help text are very thorough, almost to a fault. Their only real mistake was falling into the trap of premature optimization.
Group 13 showed a flawed understanding of key theory concepts and encountered quite a few difficulties in applying them to real world problems. They also manifest a few sparks of intuition and came up with a novel idea for the full MC approach to computing VaR. Overall, however, their report was clearly lacking key information about their work, interpretation of the results and understanding of their own mistakes.

6.3. Group 14

Group 14 exhibit an intermediate understanding of key concepts about risk management but got lost on more complex subjects. They comment their code rather well, although it was not always correct, but were lacking in comments in their paper. While they understood when they were making a mistake, they lacked any idea of why that might be.
Overall they produced a satisfactory work although not of the best quality.

6.4. Evaluation table

Our marking procedure, was to assign each point a different weight, with weights summing up 10 based our best judgement regarding each point's complexity. Moreover, we also weighted the results' correctness three times more than the coding and paper points. Finally we adjusted the marks to fit our best opinion about each group's overall work.
We evaluated our colleagues' work with the following marks:

	Group 12	Group 13	Group 14
Final marks	4	2.5	3.5

For more detail we refer back to the excel file.

7. Appendix

7.1. Closed formula for the Cliquet

We have that the payoff of the Cliquet option are the following:

$$\left[L \cdot S_{t_i} - S_{t_{i-1}}\right]^+ \qquad \text{if there is no default before } t_i$$

Hence the price at t_0 is the following:

$$Price(t_0) = \mathbb{E}^0 \left\{ \sum_{i=1}^{N} D(0,t_i) \cdot \left[L \cdot S_{t_i} - S_{t_{i-1}}\right]^+ \cdot \mathbb{I}_{\tau > t_i} + \pi\, D(0,t_i)\, NPV(t_i) \cdot \left(\mathbb{I}_{\tau > t_{i-1}} - \mathbb{I}_{\tau > t_i}\right) \right\}$$

First of all we evaluate the first terms:

$$\mathbb{E}^0 \left\{ \sum_{i=1}^{N} D(0,t_i) \cdot \left[L \cdot S_{t_i} - S_{t_{i-1}}\right]^+ \cdot \mathbb{I}_{\tau > t_i} \right\}$$

by the linearity of the mean and independence of the interest rates, default of ISP and the underlying we have:

$$\sum_{i=1}^{N} B(0,t_i) \cdot P(0,t_i) \cdot \mathbb{E}_0 \left\{ \left[L \cdot S_{t_i} - S_{t_{i-1}}\right]^+ \right\}$$

for any i we can condition the expected value with respect to $\mathcal{F}_{t_{i-1}}$:

$$\mathbb{E}_0 \left\{ \left[L \cdot S_{t_i} - S_{t_{i-1}}\right]^+ \right\} = \mathbb{E}_0 \left\{ \mathbb{E}_{t_{i-1}} \left\{ \left[L \cdot S_{t_i} - S_{t_{i-1}}\right]^+ \right\} \right\}$$

we gather the value $S_{t_{i-1}}$, which is $\mathcal{F}_{t_{i-1}}$ measurable:

$$\begin{aligned}
(1) \quad \mathbb{E}_0 \left\{ \left[L \cdot S_{t_i} - S_{t_{i-1}}\right]^+ \right\} &= \mathbb{E}_0 \left\{ S_{t_{i-1}} \cdot \mathbb{E}_{t_{i-1}} \left\{ \left[L \cdot \frac{S_{t_i}}{S_{t_{i-1}}} - 1\right]^+ \right\} \right\} \\
&= \mathbb{E}_0 \left\{ \frac{B(t_0;t_{i-1})}{B(t_0;t_{i-1})} S_{t_{i-1}} \cdot \mathbb{E}_{t_{i-1}} \left\{ \left[L \cdot \frac{S_{t_i}}{S_{t_{i-1}}} - 1\right]^+ \right\} \right\} \\
\text{thanks to the martingality and independence} \quad &= S_0 \cdot \mathbb{E}_0 \left\{ \frac{1}{B(t_0;t_{i-1})} \cdot \mathbb{E}_{t_{i-1}} \left\{ \left[L \cdot \frac{S_{t_i}}{S_{t_{i-1}}} - 1\right]^+ \right\} \right\} \\
&= S_0 \cdot \mathbb{E}_0 \left\{ \frac{1}{B(t_0;t_{i-1})} \cdot \mathbb{E}_{t_{i-1}} \left\{ \left[L \cdot \frac{S_{t_i}}{S_{t_{i-1}}} - 1\right]^+ \right\} \right\} \\
&= S_0 \cdot \mathbb{E}_0 \left\{ \frac{1}{B(t_0;t_{i-1})} \cdot C\left(L\, \frac{B(t_0,t_i)}{B(t_0,t_{i-1})}, 1, \sigma\right) \right\}
\end{aligned}$$

where we have:

$$\begin{aligned}
C\left(L\, \frac{B(t_0,t_i)}{B(t_0,t_{i-1})}, 1, \sigma\right) &= L\, N\left(d_1^{(i)}\right) - B(t_0;t_{i-1},t_i)\, N\left(d_2^{(i)}\right) \\
d_1^{(i)} &= \frac{\ln\left(\frac{L}{B(t_0;t_{i-1},t_i)}\right)}{\sigma\sqrt{t_i - t_{i-1}}} + \frac{1}{2}\sigma\sqrt{t_i - t_{i-1}} \\
d_2^{(i)} &= d_1^{(i)} - \sigma\sqrt{t_i - t_{i-1}}
\end{aligned}$$

hence we can write:

$$\sum_{i=1}^{N} B(0,t_i) \cdot P(0,t_i) \cdot \mathbb{E}_0 \left\{ \left[L \cdot S_{t_i} - S_{t_{i-1}}\right]^+ \right\} = \sum_{i=1}^{N} B(0,t_i) \cdot P(0,t_i) \cdot \frac{S_0}{B(t_0;t_{i-1})} \cdot C\left(L\, \frac{B(t_0,t_{i-1})}{B(t_0,t_i)}, 1, \sigma\right)$$

Now we pass to evaluate the second term, but first we compute the expected value of the NPV as a function of time. Recall:

$$NPV(t_i) = \mathbb{E}_{t_i}\left\{\sum_{j=i}^{N} D(t_i,t_j)\cdot\left[LS_{t_j} - S_{t_{j-1}}\right]^+\right\}$$

Hence we can write its expected value as:

$$\begin{aligned}
\mathbb{E}_0[NPV(t_i)] &= \mathbb{E}_0\left\{\mathbb{E}_{t_i}\left\{\sum_{j=i}^{N} D(t_i,t_j)\cdot\left[LS_{t_j} - S_{t_{j-1}}\right]^+\right\}\right\}\\
&= \sum_{j=i}^{N} B(t_0;t_i,t_j)\cdot\mathbb{E}_0\left\{\mathbb{E}_{t_i}\left\{\left[LS_{t_j} - S_{t_{j-1}}\right]^+\right\}\right\}\\
\text{by the tower property } &= \sum_{i=j}^{N} B(t_0;t_i,t_j)\mathbb{E}_0\left\{\left[LS_{t_j} - S_{t_{j-1}}\right]^+\right\}\\
\text{now we apply the same trick again to use } \mathcal{F}_{t_{j-1}} &= \sum_{i=j}^{N} B(t_0;t_i,t_j)\mathbb{E}_0\left\{\mathbb{E}_{t_{j-1}}\left\{\left[LS_{t_j} - S_{t_{j-1}}\right]^+\right\}\right\}\\
\text{reasoning just like we did for (1) } &= \sum_{i=j}^{N} B(t_0;t_i,t_j)\frac{1}{B(t_0,t_j)}\cdot S_0\cdot C\left(L\frac{B(t_0;t_{j-1})}{B(t_0;t_j)},1,\sigma\right)
\end{aligned}$$

Hence we can finally write:

$$\begin{aligned}
NPV(t_0) = \mathbb{E}^0&\left\{\sum_{i=1}^{N} D(0,t_i)\cdot\left[L\cdot S_{t_i} - S_{t_{i-1}}\right]^+\cdot\mathbb{I}_{\tau>t_i} + \pi\, D(0,t_i)\, NPV(t_i)\cdot\left(\mathbb{I}_{\tau>t_{i-1}} - \mathbb{I}_{\tau>t_i}\right)\right\}\\
&= \sum_{i=1}^{N}\left\{B(t_0;t_{i-1},t_i)\cdot S_0\cdot P(t_0,t_i)\cdot C\left(L\frac{B(t_0,t_{i-1})}{B(t_0,t_i)},1,\sigma\right)+\right.\\
&\left.+\ \pi\, B(t_0,t_i)\cdot[P(t_0,t_{i-1}) - P(t_0,t_i)]\cdot S_0\cdot\left[\sum_{j=i}^{N}\frac{B(t_0;t_i,t_j)}{B(t_0,t_j)}C\left(L\frac{B(t_0;t_{j-1})}{B(t_0;t_j)},1,\sigma\right)\right]\right\}
\end{aligned}$$

This ultimately yields the closed formula for the Cliquet option:

$$\begin{aligned}
Price(t_0) = \sum_{i=1}^{N}&\left\{B(t_0;t_{i-1},t_i)\cdot S_0\cdot P(t_0,t_i)\cdot C\left(L\frac{B(t_0,t_{i-1})}{B(t_0,t_i)},1,\sigma\right)+\right.\\
&\left.+\ \pi\ \cdot[P(t_0,t_{i-1}) - P(t_0,t_i)]\cdot S_0\cdot\left[\sum_{j=i}^{N} C\left(L\frac{B(t_0;t_{j-1})}{B(t_0;t_j)},1,\sigma\right)\right]\right\}
\end{aligned}$$

7.2. About the use of Logarithmic returns

Logarithmic returns are widely used in finance because they provide an accurate change in the value of an asset over a period of time. The key result that can be exploited is the fact that log returns are additive: the return of a portfolio composed by different assets is simply the sum of the logarithmic returns. Moreover, considering our scenario in which a generic asset follows a geometric Brownian motion it is clear that by considering Logarithmic returns it is possible to work in a environment where variables are considered to by normal distributed. Let's now take a deeper dive into a number of powerful arguments to justify the use of logarithmic returns:

- Logarithmic returns can be interpreted as continuously compounded returns: when considering non stochastic processes the frequency of compounding does not matter and returns across assets can be more easily compared
- Continuously compounded returns are time additive and it is easier to derive the time series properties of additive processes than multiplicative processes.
- The use of logarithmic returns prevents security prices from becoming negative in models of security returns.
- If a security price follows geometric Brownian motion the logarithmic returns are normally distributed

8. References

- Principal Component Analysis For Stock Portfolio Management, by Giorgia Pasini, University of Verona
- A Comparison between Logarithmic and Simple Returns, by Robert S Hudson, Newcastle University Business School
- "Arbitrage Theory in Continuous Time" T.Bjork
- On the Pricing of Cliquet Options with Global Floor and Cap, by Mats Kjaer, Department of Mathematics Chalmers University of Technology and G¨oteborg University