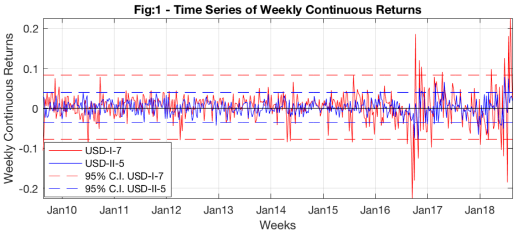
Descriptive Statistics Stock prices are generally assumed to follow a geometric Brownian motion $dS = \mu S dt + \sigma S dz$, where μ is the expected return, σ is the volatility and dz is a Wiener process. This model implies that continuous returns are normally distributed as $ln(S_T/S_0)/T \sim \phi(\mu - \sigma^2/2, \sigma^2/T)$. This allows to establish a precise relationship between prices and returns, where the first two moments of the price distribution depend on the expectation and the variance of the returns. For example, although the allocation of risk capital is based on the value of a portfolio, the returns dynamics still impact the risk capital through the price distribution. Even though the Jarque-Bera test rejects the null hypothesis that continuous returns come from a normal distribution at the 5% significance level, a plot of the 95% confidence interval (Fig:1) and a graphical analysis of the distributions allow to conclude that this assumption is not too far off, especially for the second asset.



The additivity of continuous returns in the time series allows to easily compute the expected weekly returns of the two assets ($\bar{r}_1 = 0.27\%; \bar{r}_2 = 0.18\%$), the variances ($\sigma_1^2 = 0.17\%; \sigma_2^2 = 0.04\%$) and the standard deviations ($\sigma_1 = 4.09\%$; $\sigma_2 = 1.92\%$), which restore the original measurement unit. These statistics confirm what is visible from the time series plot of returns, i.e. that the two securities have continuous returns which fluctuate around a mean very close to zero, and that higher expected return is associated with higher volatility. This is a clear financial relationship, according to which instruments that provide a higher expected return are generally also associated with a higher risk: what matters is the not return, but the risk adjusted return.

While kurtosis ($k_1 = 9.6975$; $k_2 = 5.0236$) is a measure of the weight in the tails of the distribution (or of its peakedness), skewness ($s_1 = -0.2790$; $s_2 = -0.1423$) is a measure of its (a)symmetry around the mean. The normal distribution is known to have a population kurtosis of 3 and a skewness of 0. Therefore, the two assets under study are leptokurtic (or with a relatively peaked distribution) and have longer (i.e. with more mass) left-tails. This means that, for both stocks, and especially for the first one, extreme positive or negative returns are more likely than under normality (a feature which "boosts" the volatility) and that, historically, there have been frequent small gains and a few extreme losses. Although this last feature is not economically very significant for both securities, it matters for an investor, as it indicates a greater chance of extremely negative outcomes, and results in the mode being greater than the median, which in turn is greater than the mean.

The covariance ($cov_{12} = 0.03\%$) measures the linear association between the two series of continuous return, representing their joint variability. The positive value indicates that high returns for the first stock tend to be associated with high returns for the second stock. However, the magnitude of this value denotes a weak linear relationship.

Although the correlation ($\rho_{12} = 34.69\%$) is nothing more than a standardization of the covariance by the volatilities of the two assets, the former overcomes one important weakness of the latter, namely that it does not have an upper and lower bound and that it is greatly influenced by the scaling of the numbers. Moreover, it is important to notice that the positive correlation coefficient does not imply any relationship of causality between the two variables, as it merely describes the strength and the direction of the linear relationship. In this case, the correlation between the two securities is positive, though off from a perfect relationship, indicating that when the return of one stock is high, then the return of the other has a higher probability (than under independence) of being high. It is clear that the correlation is more useful to interpret the relationship between the two stocks than the covariance. To obviate some of the weaknesses of the sample correlation coefficient (such as that it can be strongly influenced by outliers, that it cannot identify non-linear relationships and that tests based on it rely on the assumption of normality) it is useful to compute the Spearman coefficient ($r_s = 17.86\%$, significantly different from zero at a p-value=0.0001), which is a nonparametric measure of rank correlation. This measure does not require any assumption of linearity and captures monotonic relationships (based on ranks) between the returns of the two assets: the positive measure obtained indicates an imperfect positive monotonic dependence, meaning that when the return of one stock is increases, the return of the other tends to increase. Finally, the fact that Pearson's measure is twice as big as Spearman's raises some questions about the meaningfulness of the former: indeed, by plotting a scatter plot it is easy to identify the presence of outliers.

There are other useful statistics to describe the distributions of the returns. The one week 97.5% VaR is the threshold value such that the probability of a greater than it over the week is 2.5%. It can be computed with three different approaches: historical ($VaR_1 = -9.11\%; VaR_2 = -3.97\%$), by rounding down the 2.5th percentile of the historical distribution; variance-covariance ($VaR_1 = -7.75\%; VaR_2 = -3.59\%$), by assuming normal returns; Monte Carlo ($VaR_1 = -7.84\%; VaR_2 = -3.63\%$), by applying the historical approach to a simulated distribution. Comparing the values obtained for different confidence levels, the last two approaches give results close to the first one at the 5% significance level, but their difference widens at the 1% level: this is further indication of the leptokurtosis of the actual distributions. Finally, the shortfall expectation ($\bar{r}_{1l} = -13.42\%; \bar{r}_{2l} = -5.40\%$) and shortfall volatility ($\sigma_{1l} = 4.58\%; \sigma_{2l} = 0.64\%$) complement the VaR by informing about the tail distribution.

The transformation $f(x) = 100\, e^x$ shows how an investment of 100 would evolve in one week's time. Alternatively, it can be interpreted as giving a form of weekly discrete return normalized around 100 (or in percentage). It is then interesting to notice that simple and continuous returns yield different results, with the latter always underestimating the former, when the ratio of the consecutive prices is far from 1. The covariance ($cov_{12} = 2.7261$) and the variances ($\sigma_1^2 = 16.7871; \sigma_2^2 = 3.7024$) display their pre-announced sensitivity to the magnitude of the variables, which are now much bigger. Moreover, their values do not increase exactly by the same proportion due to the non-linearity of the transformation. For the same reason, the Pearson coefficient ($\rho_{12} = 34.58\%$) is not exactly unchanged in absolute terms, but it is in the magnitude, as it scales for the standard deviations ($\sigma_1 = 4.0972; \sigma_2 = 1.9242$). Conversely, Spearman's measure is exactly unchanged and still significant. Indeed, as explained before, this measure is based on ranks, rather than linearity, and is therefore unaffected by a non-linear monotonous transformation. Regardless of the way return is computed, both the log return and its monotonous transformation still represent the same changes in the prices of the variables and the rank is therefore unchanged.

Descriptive Statistics for Quartile Ranges Computing the quartiles of the first asset allows to partition the two-dimensional time series of log returns into four quartile ranges. Repeating the same procedure also for the second asset, the following statistics can be identified.

Var-Cov		Q1		Q2		Q3		Q4	
		I-7	II-5	I-7	II-5	I-7	II-5	I-7	II-5
wrt I-7	I-7	0.12%	0.03%	0.00%	0.00%	0.00%	0.00%	0.10%	0.03%
	II-5	0.03%	0.06%	0.00%	0.02%	0.00%	0.02%	0.03%	0.05%
	Cor	37.51%		-6.14%		-5.76%		40.20%	
wrt II-5	I-7	0.24%	0.03%	0.10%	0.00%	0.11%	0.00%	0.19%	0.02%
	II-5	0.03%	0.02%	0.00%	0.00%	0.00%	0.00%	0.02%	0.01%
	Cor	39.79%		6.36%		7.83%		40.58%	

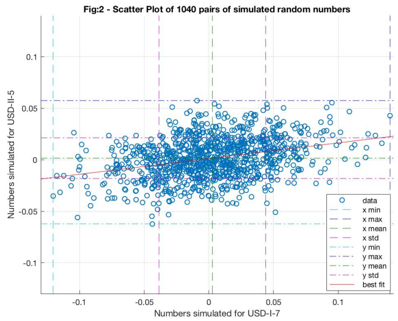
Independently of whether the quartiles are computed with respect to the first or the second asset, variances, covariances and correlations are greatest in Q1 and Q4 and the latter are also more similar across "wrt I-7" and "wrt II-5". This is explained by the fact that Q1 and Q4 contain the most extreme positive and negative returns (which increase volatility) and that the two assets co-move more closely under extreme scenarios. This analysis can be interpreted in a context of regime switching. When volatility is high, for example when we are in a bull (e.g. tax deregulation) or bear market (e.g. stock market crash), the returns of the two securities will be more correlated as they will both reach uncommonly high or low returns. On the other hand, when the market is affected by low volatility, the returns of the two assets will have a weaker linear relationship, as they will fluctuate noisily around their means.

Cholesky Decomposition The variance-covariance matrix Σ is a positive-semidefinite matrix, as it is a symmetric n x n real matrix for which the scalar $W'\Sigma W$ (where W is the weighting vector) is non-negative for every non-zero column vector W of n real numbers. Geometrically, this can be interpreted as a quadratic function which has a non-negative curvature in every direction W. Indeed, independently of the vector of weights W, the scalar $W'\Sigma W$, which represents the variance of a portfolio, is always positive and it is the expression of the curvature of the function along the direction W. Note also that the value of the curvature is constant, as the second derivative of a quadratic function is constant. In the two asset case, the covariance matrix is positive semidefinite because the values along the main diagonal (i.e. the variances of the two assets) and the determinant are non-negative. The power of a positive semidefinite matrix is that it allows to implement the Cholesky decomposition. Effectively, this factorization consists in taking the square root of the variance-covariance matrix and generate a lower triangular matrix (i.e. an asymmetric matrix with zero coefficients above the diagonal), such that the product of this matrix by its transpose is the variance-covariance matrix itself ($DD' = \Sigma$). The lower triangular matrix generated is the following:

$$D = \begin{bmatrix} 0.0409 & 0.000 \\ 0.0067 & 0.0180 \end{bmatrix}$$

Random Numbers Stochastic simulation is an easy tool to reproduce realistic market dynamics and generate distribution functions, which enable a portfolio manager to determine the risk exposure of its holdings, and therefore the required risk capital. For this purpose, the Cholesky decomposition can be used to simulate a bi-variate distribution for $X = \binom{x_1}{x_2} \sim N(\xi, \Sigma)$, where ξ is the matrix of expected weekly returns and Σ is the non-singular variance-covariance matrix, as calculated from the historical time series. It is sufficient to generate a matrix (nx2) of stochastically independent (and therefore uncorrelated), uniformly [0,1] distributed random numbers, make them standard normally distributed (for example by applying the direct method), multiply the vector by the triangular matrix (2x2) generated through the Cholesky decomposition and add the expected values (nx2): $R \sim U[0,1] \rightarrow Z \sim N(0,1) \rightarrow X = \xi + ZD'$

The scatter diagram (Fig:2) plots 1040 pairs of simulated random numbers with the above mentioned distribution and helps to investigate the properties of the two series and their relationship.

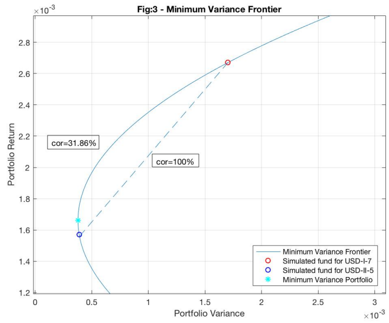


Although the multivariate normal distribution has been generated to have a specific return and variance-covariance matrix, its moments are not exactly the desired ones, due to the generation of numbers which are not perfectly standard normal. The point cloud is clustered around the two positive means ($\bar{r}_{s1} = 0.27\%; \bar{r}_{s2} = 0.16\%$), which are respectively equal and slightly lower than the target ones (to the precision level of one basis point). The series simulated for the first asset displays a higher variation, as shows by the standard deviations ($\sigma_{s1} = 4.13\%;$ $\sigma_{s2} = 1.97\%$) and the ranges ($range_{s1} = 0.2606; range_{s1} = 0.1197$). This is again in line with the positive financial relationship between expected reward and risk. When looking at the variance-covariance matrix ($\sigma^2_{s1} = 0.17\%; \sigma^2_{s2} = 0.04\%; cov_{s12} = 0.03\%$), the values are in line with the target ones (again, only to the precision level of one basis point) and allow to draw the same conclusion about the positive linear association between the returns of the two assets. The correlation coefficient ($\rho_{s12} = 31.86\%$), which is lower than the historical one, provides a standardized measure of this linear relationship. The scatter plot is useful to graphically identify the direction of the relationship between the series of simulated returns, as inferred from the slope of the best fit line, and its strength, weak in this case, as shown by the fact that the simulated data points are noisily clustered around the best fit line. What really changes are the values of kurtosis ($k_{s1} = 3.1967$, $k_{s2} = 2.8707$) and skewness ($s_{s1} = 0.0454$, $s_{s2} = 0.0074$), which, by construction, now resemble more closely those of a normal distribution and therefore underestimate the actual tail risk. Indeed, the values of the fourth moments are now closer to 3 and indicate slightly leptokurtic and platykurtic distributions respectively, in contrast with the heavy tails of the historical series. This means that, while for the first simulated series tail events are slightly more likely than under the normal distribution, for the second the opposite holds. The values of the third moments are now closer to 0 and indicate very slightly positively skewed distributions, in contrast with the negative skewness of the historical series, which indicate the tendency to frequent small losses and a few extreme gains.

The basic mean-variance optimization is a non-linear problem which consists in minimizing the variance of the portfolio ($min\ W'\Sigma W$), while targeting a specific expected return ($\bar{R}'W = \mu$), by modifying the vector of weights ($e'W = 1$). In general, a convex optimization problem is defined by the minimization (maximization) of a convex (concave) function over a convex decision space. Since it minimizes a convex function with linear constraints, this is a convex optimization problem, which implies the sufficiency (and necessity) of the first order condition: this means that any local solution is also a global solution or, in other words, that any solution stating that the first derivative is zero is the global minimum, i.e. the optimum. The function which represents the variance of the portfolio $W'\Sigma W$ is convex because (this is a necessary and sufficient condition) its Hesse matrix is positive semi-definite, as Σ itself is positive semi-definite. Indeed, the second derivative of a quadratic function, which measure the curvature of the function along the direction W, is constant (2Σ). This means that the curvature stays the same, independent of w_i, as long as you fix the direction W along which the second derivative is taken.

The Lagrange approach provides a solution to this convex optimization problem, by minimizing a function which is a combination of the original objective function and of its constraints. This increases the dimension of the optimization problem but allows to reduce it to an unconstrained case. Without going into further detail, it is important to understand that this transformation is responsible for the saddle structure of the Lagrange function and for the fact that, at the optimum, the gradient of the original objective function is a linear combination of the gradients of the constraint functions. The Lagrange approach has a strict economic interpretation in that the Lagrange multiplier can be seen as the shadow price of the i^{th} restriction, i.e. the price that one would pay for an additional unit of the i^{th} resource. Moreover, as the Lagrange multipliers are linearly dependent on the target return μ and as the optimal value of the vector of weights W is linearly dependent on the Lagrange multipliers, it follows that W is linear in μ as well. This means that we have a linear function of portfolio weights (the solution W^* of the Lagrange problem can also be written as the weighted sum of two vectors W^1 and W^2 where $e'W^i = 1$ for 1=1,2) and all that is necessary are two efficient portfolio weights, because then every linear combination is efficient. This implies that optimal portfolios must be convex combinations of portfolios W^1 and W^2, which are independent from investor's μ (two fund theorem). What is responsible for the separation theorem is the fact that $W'\Sigma W$ is a quadratic function and, given linear equalities as constraints, the Lagrangian becomes a quadratic function as well, with linear F.O.C. Finally, it is important to notice that the introduction of a non-linear constraint (such as a restriction on short selling or a limit on the exposure to foreign currency) would invalidate the two fund theorem.

The weights obtained for the minimum variance portfolio are $w_1^* = 8.26\%$ and $w_2^* = 91.74\%$. The discrepancy in the allocation is due to the high volatility of asset 1, which has a much worse return-variance ratio. The minimum variance portfolio is expected to have a return $r_{mvp} = 0.17\%$ and a variance $\sigma^2_{mvp} = 0.04\%$. The portfolio variance is lower than that of the individual assets due to the benefit of diversification, which appears whenever two assets are less than perfectly positively correlated. Indeed, as the expected return is not affected by the correlation and as the lower the correlation the greater the gain in efficiency, in mean variance optimization, portfolios with low correlation are preferred. The following graph (Fig:3) plots the minimum variance portfolios.

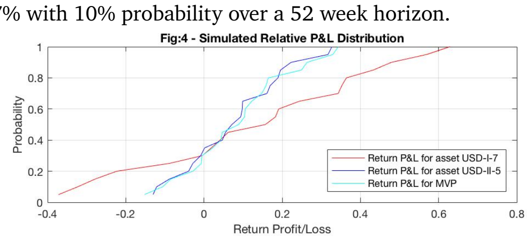


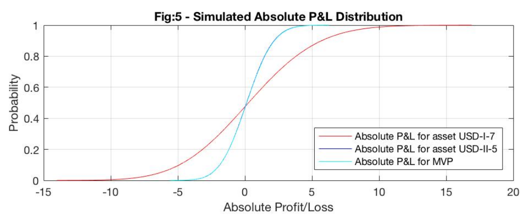
The solid line shows the portfolio opportunity set for $\rho_{s12} = 31.86\%$, i.e. all the combinations of portfolio expected return-variance that can be constructed from the two available assets. The structural property of the minimum variance frontier is a parabola in the σ^2-μ space, because W is linear in μ and the square of a linear function ($W'\Sigma W$) is quadratic. Moreover, the upper part strictly dominates the lower one and, in this section, the slope is decreasing as the marginal rate of substitution between risk and return declines. Finally, the dashed line shows that, as mentioned earlier and contrary to the case studied, a perfect linear correlation does not allow to benefit from diversification.

The 95% confidence interval for the expected weekly return of the minimum variance portfolio is $[0.05\%; 0.28\%]$, while the 95% prediction interval for the return realization is $[-3.65\%; 3.99\%]$. As expected, the latter is wider than the former, as it is more difficult to estimate a range within which an individual value (the next realization) will fall, rather than doing so for the average value. In other words, the range of errors or variance is larger when forecasting a single value compared to forecasting the mean. Both have the same interpretation: there is a 95% probability that this interval includes the true prediction or the true mean of the return.

Simulation By breaking down the simulated series X into 20 paths, each covering a period of 52 weeks, it is possible to implement a P&L simulation for a static portfolio, which is invested with the minimum variance weights in two assets ($P_1 = 100, P_2 = 80$ at the beginning of the holding period). However, this optimization is not ideal, as rebalancing the portfolio would allow to reallocate the resources according to the regime switching and the changing volatilities. Indeed, the dependence on the volatility structure changes over time and a dynamic allocation would be concerned with the density distribution of returns and prices.

The plot of the relative P&L distributions (Fig:4) shows the P&L in terms of returns obtained over a 52-week period for 20 simulated paths. It shows how the MVP portfolio tracks more closely the second asset, according to the weights assigned, and that there will be a loss after 52 weeks with a probability of 30%. An alternative representation is the plot of the absolute P&L distribution (Fig:5), which shows the P&L in terms of absolute gain/loss. For this plot, it was assumed that the continuous returns (prices) come from an exact normal (log-normal) distribution, with the preferred parameters. It shows how the distribution of the MVP is steeper than the one of the first asset and therefore it is less risky. It also shows how the area between the cyan and red curve above the point where they cross is bigger than the one below, indicating that the upside potential is larger than the downside. Moreover, this plot allows to understand why the $\mu - \sigma^2$ space is not sufficient to conduct a sound analysis. Indeed, the distributions are particularly useful in the context of risk management, as they allow to compute the shortfall risk, i.e. the probability of obtaining a loss grater than a given value. For example, a pension fund, which is faced with the risk of not being able to cover its liabilities, and which normally targets a return associated with a specific level of risk, might be interested in knowing that this fund will lead to a loss greater than 10.57% with 10% probability over a 52 week horizon.





Note: all calculations and plots were done in MATLAB.