University College London

Department of Computer Science

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science in Computational Finance, University College London

Currency hedging in iVaR setting.

Author
Jannis Fierens

Academic Supervisor
Prof. Germano Guido
Department of Computer Science
University College London

Industrial Supervisor
Mr. Lemahieu Emiel
Dr. Wyns Maarten
Dr. Achtsis Nico
Quantitative Analytics
InvestSuite

September 26, 2022

This dissertation is submitted as part requirement for the MSc Computational Finance degree at UCL. It is substantially the result of my own work except where explicitly indicated in the text.

Abstract

This paper makes a humble attempt to put currency hedging in a more realistic setting by considering different hedging approaches. Continuously one-to-one hedging is an extremely costly procedure which is why retail investor are advised to construct an optimal basket of currencies for a predetermined period. The literature has recently been focusing on minimising either the minimum or the mean-variance of an international hedged portfolio with a basket of foreign currencies. Other risk measures applied to hedging are Value at Risk and Tail Value at Risk. One could argue that these measures are not capable of entirely capturing what an retail investor perceives as risk. Currency hedging is therefore put in a brand new perspective by considering the state of the art iVaR risk measure. It was introduced by the Belgian financial technology company InvestSuite and embraces the retail perception of risk in a portfolio: the frequency of drawdowns, their magnitude, and the time to recover from them. A rather ambiguous concept in the literature is the definition of hedging applied to the foreign currency market from the perspective of a retail investor. The realistic problem setting inhibits strict bounds for the asset allocation which implies that some methodologies might not be as optimal anymore from a risk perspective. In particular, the iVaR optimisation problem is compared to a static benchmark and a minimum variance portfolio. This paper does consider different estimates for the covariance matrix. These methods are compared on their out of sample realised volatility, hedging effectiveness, Sharpe ratio, Calmar and pain ratio, cumulative drawdown, average drawdown and average drawdown reduction. Finally, Hansen's bootstrapped model confidence set is employed to determine if the difference is also statistical significant. This paper finds sufficient evidence that results obtained from constructing a dynamic hedging strategy in a strict setting are not statistically different. These findings should not be attributed to an ill definition or the appropriateness of different methodologies, but rather hinges to the problem setting. For instance, retail investors that seek to mitigate their risk ought not to increase the FX exposure above the level of their initial investment in order to gain diversification benefits. Furthermore, the FX market is characterised by co-movement and clusters of shared risk. Once these strict bounds are relaxed, the methodologies will gain in their effectiveness.

Acknowledgments

First and foremost, I would like to thank Professor Guido Germano for meeting me on a weekly basis and giving me the flexibility to conduct my research at InvestSuite. Extremely helpful was Emiel during my research, since he was always available when help was needed and willing to go above and beyond. I am also grateful for the sharp insights of Maarten and Nico who clearly set the scene of this study. Lastly, I am particularly grateful to my parents and loved ones for encouraging me during this process.

Contents

Chapter 1

Introduction

Since the Bretton Woods collapsed in 1973, investors that hold an international portfolio of financial instruments bear additional risk of trading foreign currencies for their national currency. The foreign exchange market (FOREX, FX) is the largest financial market these days and therefore also extremely liquid. The bulk of these transactions occur only in eight currencies, i.e. U.S. the dollar (USD), the EURO (EUR), the British pound (GBP), the Japanese yen (JPY), the Swiss franc (CHF), the New Zealand dollar (NZD), the Australian dollar (AUD) and the Canadian dollar (CAD), while approximately 182 circulate around the world (Petropoulos et al. (2017)). This significantly facilitates trading which appeals to speculants and arbitrageurs (Ozturk et al. (2016)). Nevertheless, retail investors aim to have a smooth ride and business strive to hedge FX risk. One of the main drivers of these FX time series is a country's monetary policy. Central banks attempt to maintain price stability by influencing the money supply, interest rates and inflation. Beine et al. (2006); Égert and Kočenda (2014), amongst others, found empirically that FX time series are materially influenced by monetary, geopolitical and economic factors and therefore serve as an overall economic activity aggregator. This real exchange risk ought to be avoided by risk minimising agents. The continuously hedging process is however costly and sometimes even not feasibe. Therefore, according to the investor's risk tolerance, it is optimal to always hold a basket of foreign currencies. The question then remains: How do investors perceive risk such that they can determine their optimal basket of currencies?

Investors in general strive to have monotonic growth such that their time in the market is maximised and they are not forced to liquidate their portfolio because they could not stomach the volatility. Traditional approaches for measuring risk turn out to be not perfectly aligned to what investors perceive as risk. For instance, risk is seen as both the frequency and severity at which losses occur and therefore are asymmetric. Using the returns' variance would be inappropriate since investors might not care or even seek positive volatility. Value at Risk (VaR) and Expected Shortfall (ES) are asymmetric but fail to incorporate the time dimension, that is the total period a portfolio is under water, that is the loss compared to the previous high water mark. Retail investors have not only the tendency to compare the best and worst absolute levels of their assets, but also consider the time they are (not) at these levels. The most natural risk measure to model investors' behaviour is therefore the integrated time value at risk (iVaR) which takes these three dimensions into account. In other words, the optimal portfolio is constructed in a way that minimises the average drawdown compared to the running maximum.

The contribution of this paper to the academical literature is twofold. It is the first time that an iVaR objective has been applied to currency hedging. Second, applying different optimisation methodologies when the problem exhibits strict bounds seems not to give any significantly different results in the expected path of the basket. The remainder of the paper is structured as follows: Both the history of different risk measures and how currency hedging is described in literature are presented in chapter 2. Next, chapter **??** elaborates on the methodology used for calculating an optimal iVaR portfolio, estimating a regularised (Ridge and Lasso) variance-covariance matrix and the calibration of both th constant and dynamical correlation coefficient (CCC and DCC) generalised autoregressive conditional heteroskedastic in order to predict the variance-covariance matrix. This chapter also describes the model confidence set such that we statistically can compare these different methodologies according to the chosen metric. The data description and results are given in chapter **??**. Finally, the concluding findings of this paper are summarised in chapter **??**.

Chapter 2

Literature review

2.1 Portfolio optimisation: an evolution in risk measures

2.1.1 Variance as a risk measure

Asset managers and robo-advisors today still make use of Markowitz' mean-variance portfolio optimisation framework. Harry Markowitz introduced a mathematical methodology and algorithm Markowitz (1952, 1959) for the optimal allocation of assets under uncertainty based on the second central moment μ_2, i.e. using the variance as a risk measure

$$\sigma_p^2 = \mu_{p,2} = \int_{\mathbb{R}} (x - \mu)^2 dF_{X_p}(x) \tag{2.1}$$

where X_p is the portfolio's return which is assumed to be normally distributed with mean $\mu = \mathbb{E}[X_p]$, also known as the first raw moment, the variance σ_p^2 and $dF_{X_p}(x)$ the normal probability function. Today, this theory is famously known as modern portfolio theory (MPT) which lays down the optimal trade off between expected return and variance also referred to as the mean-variance tradeoff. It has a wide range of applications in financial theories (e.g. Capital Asset Pricing and Efficient Market Hypothesis) and mathematical finance (e.g. in option pricing where assets are considered to follow a geometric brownian motion and where variance if one of the most important parameters Black and Scholes (1973)).

MPT assumes that investors have access to all financial instruments in the entire world and therefore can optimally choose the weights $\boldsymbol{w}$ assigned to each instrument such that the ratio expected return over risk (i.e. portfolio's variance) is optimal which is also known as the Sharpe ration. For instance, a portfolio that has the same expected return but a higher variance is not optimal. As a result, the efficient frontier could be obtained by calculating the best ratio return versus risk for every possible expected return. Since returns are assumed to be normally distributed, this objective function is always convex. Finally, the straight line starting from the risk-free return and tangent to the efficient frontier is maximises the Sharpe ratio and composes all combinations of the risk free return and the optimal risk portfolio.

Besides the mathematical framework, Markowitz also introduced the Critical Line Algorithm (CLA) which solves this quadratic optimisation problem. This optimisation method in fact solves a series of different convex problems and attempts to improve the Sharpe ratio by adding each time a financial instrument. In doing so, the narrative of considering an overall risk-adjusted portfolio and not merely the sum of risky financial instruments, got established. Markowitz (1952) was one of the most influential papers to consider a quantitative approach to the investment problem.

Although MPT has many appealing properties, both quantitative and qualitative, it lacks realistic assumptions. The mean-variance framework for instance implicitly assumes that the assets' variances are all finite and constant throughout time. The true variance-covariance matrix $\mathbf{\Sigma}$ has to be full-rank as well in order to calculate the optimal portfolio. This means that in this financial market we cannot replicate an asset's time series by dynamically trading the other securities. Furthermore, assets have positive variances which implies that $\mathbf{\Sigma}$ is positive definite. Another drawback of the mean-variance framework, is the assumption of the existence of an unbiased estimate of the expected return. This thesis therefore only considers obtaining the global minimum of the portfolio's volatility σ_p as objective function by choosing the optimal weights $\mathbf{w}^* \in \mathbf{\Theta}$ with $\mathbf{\Theta}$ the set of all possible combinations for N assets $\mathbf{\Theta} \subset \mathbb{R}^N$. Given the true variance-covariance matrix $\mathbf{\Sigma}$, the convex optimisation problem is formulated as

$$\begin{aligned} &\underset{\mathbf{w} \in \mathbf{\Theta}}{\arg\min}\, \mathbf{w}^T \mathbf{\Sigma} \mathbf{w}, \\ &\text{s.t. } \mathbf{w}^{\mathsf{T}} \mathbf{1} = 1. \end{aligned} \tag{2.2}$$

The used risk measure implicitly assumes that returns are normally distributed while these are most likely to be leptokurtic distributed, i.e. a kurtosis higher than three. Value at risk and expected shortfall try to address these shortcomings by only considering the left tail of the distribution. In addition, variance is symmetric and sign-ignorant whereas investors most likely care less about upside volatility compared to downside volatility and shouldn't punish it in the same manner.

2.1.2 Semivariance as risk measure

In the same period Roy (1952) defined semi-variance as the risk measure that should be used in order to make an optimal allocation between financial instruments. Markowitz even recognised the importance of Roy's work by stating that if he published his work earlier, the optimal set of mean-variance porfolios would have been called the set of mean-semi-variance portfolios (see Nawrocki (1999)). This risk measure is formally defined as

$$\mathrm{SV}_p(y) = \int_{\mathbb{R}} \{(x-y)^2\}_+ dF_{X_p}(x) \tag{2.3}$$

where $y \in \mathbb{R}$ is the threshold at which investors start to care about potential losses. A well known variant of this risk measure is the lower half of the variance, i.e. substituting y for μ

$$\mathrm{SV}_p(\mu, \alpha) = \int_{\mathbb{R}} \{(x-\mu)^2\}_+ dF_{X_p}(x). \tag{2.4}$$

Notice that for symmetrical probability distributions the relationship $\sigma_p^2 = 2\,\mathrm{SV}_p(\mu)$ holds. The risk measure $\mathrm{SV}_p(\mu)$ could also be interpreted as a portfolio's skewness measure since the latter relationship won't hold for skewed underlying distributions.

2.1.3 Lower partial moments as risk measure

While the semi-variance risk measure is a generalisation of the second central moment, by including skewness, C Fishburn (1977) took it one step further by defining lower partial moments (LPM) in 1970

$$\mathrm{LPM}_p(\mu) = \int_{\mathbb{R}} \{(x-\mu)^\alpha\}_+ dF_{X_p}(x). \tag{2.5}$$

where $\alpha \in \mathbb{R}$. Notice that alpha is now is a hyperparameter compared to squaring the difference. By adding this flexibility, some of the issues that were accompanied by previous loss risk measures are resolved.

2.2 Currency Hedging

Hedging activities typically take place in the futures market where two parties come to an agreement to buy and sell a given amount of the underlying at an agreed upon certain date in the future, at an agreed upon price, and at a given location. This market brings predominantly hedgers, speculators and arbitrageurs together. Hedgers are typically producers, consumers or merely have the ownership of the underlying and wish to get mitigate the borne price risk. Arbitrageurs and speculants, on the other hand, want to exploit mis-pricings in this market and try to gain alpha returns. Those two groups are therefore fundamentally different in nature, hedgers are risk minimisers while speculants and arbitrageurs are willing to take on additional risk to maximise profits.

In case that markets are efficient, futures prices have to be the unbiased estimator of spot prices at the maturity of the contract which resembles to the cointegration of the current futures prices and the future spot prices. For a complete overvies, this paper refers to Dwyer Jr and Wallace (1992); Chowdhury (1991); Crowder and Hamed (1993). The literature is extensive in finding answers where exactly price discovery takes place. For an overview, please consult .

Since this paper concerns more about the inherent long-term nature of the asset class compared to the intraday movements in the foreign currency market, the focus will be on finding a basket of currencies that minimise certain risk measures. Tong (1996) showed prices in these two markets are closely tied by the arbitrage forces. The difference between the futures and spot market is therefore neglected and only spot market data is used. The next three sections elaborate on more general hedging approaches in both the univariate and single-variate setting as well as a measure the determine the hedging effectiveness.

2.2.1 Univariate hedging

In case an investor is interested in buying a financial instrument which is denominated in a foreign currency but wishes to minimise this FX, he might continuously be buying and selling futures contracts. This however could turn out to be a costly process. Another possibility is to periodically minimise the aggregate portfolio's risk $Y_t = X_t - \delta_t Y_t$, where δ_t is the hedge ratio and $(X_t)_{t\in\mathbb{N}}$, $(Y_t)_{t\in\mathbb{N}}$ are respectively the return's discrete stochastic processes of a financial asset and the foreign currency's futures contract with $t \in \{0, 1, ..., T-1, T\}$. This obviously raises the question: What type of risk are we willing to minimise for? In general the investor wants to minimise the fluctuations that are borne form the additional FX risk. Johnson (1976) expresses the conditional variance of this portfolio as

$$\mathrm{Var}(Z_t|\mathcal{F}_{t-1}) = \mathrm{Var}(X_t|\mathcal{F}_{t-1}) - 2\delta_t \mathrm{Cov}(X_{1,t}, Y_t|\mathcal{F}_{t-1}) - \delta_t^2 \mathrm{Var}(Y_t|\mathcal{F}_{t-1}) \tag{2.6}$$

where $\mathcal{F}_{t-1}$ is the filtration available up and until time $t-1$. The formal definition of the sigma-algebra is given in section (3.4.1). In order to determine the optimal hedge ratio (OHR), we calculate the partial derivatives with respect to δ_t and subsequently put the equation equal to zero. After rearranging, we obtain

$$\delta_t|\mathcal{F}_{t-1} = \frac{\mathrm{Cov}(X_{1,t}, X_{2,t}|\mathcal{F}_{t-1})}{\mathrm{Var}(X_{2,t}|\mathcal{F}_{t-1})}. \tag{2.7}$$

In case the optimal hedge ratio would not be dependent on t, it would become a static hedge. Nevertheless, financial markets typically inhibit volatility clusters and changing correlations structures among random variables. Multivariate distributions will change throughout time by the arrival of new information and a static hedge will be inadequate in reducing the portfolio's risk. The literature shows that hedge ratios are typically not stationary in time. For an overview, see Baillie and Myers (1991); Cecchetti et al. (1988); Myers (1991); Kroner and Sultan (1993); Choudhry (2003); Park and Switzer (1995).

2.2.2 Multi-variate hedging

For a while the literature focused only on single hedges, while investors hold large portfolios of different type of assets denominated in various currencies. Optimally choosing the position taken in each currency to reduce risk of the aggregate portfolio is paramount. As Markowitz pointed, we cannot isolate assets from each other and therefore ignore the covariance structure. Estimation procedures which measure hedge ratios for each hedging

instrument in isolation tend to over-estimate the number of futures contracts required to hedge the cash position and lead to a suboptimal mix in the composition of the hedge portfolio. Gagnon et al. (1998) consider a situation where an N-currency portfolio is hedged with N-futures contracts. This study is extremely relevant for this research which is why we borrow some of the equations therein. In the most general setting, we have a space of N log returns of the spot currency market $(\mathbf{X}_t)_{t\in\mathbb{N}} = (X_{1,t}, ..., X_{N,t})_{t\in\mathbb{N}}$ with respect to the domestic currency and $(\mathbf{Y}_t)_{t\in\mathbb{N}} = (Y_{1,t}, ..., Y_{N,t})_{t\in\mathbb{N}}$ the futures' log returns. It is assumed that every currency pair on the spot has also a liquid futures market. Given a multi-period setting, it is possible to express the hedged portfolio's log return by

$$Z_t = \mathbf{X}_t^\mathsf{T} \boldsymbol{\gamma}_{t-1} - \mathbf{Y}_t^\mathsf{T} \boldsymbol{\delta}_{t-1} \tag{2.8}$$

where the observation period is discrete $t = 1, ..., T$ and $\boldsymbol{\gamma}_{t-1}$, $\boldsymbol{\delta}_{t-1}$ are respectively the $N \times 1$ vector of spot and futures currency holdings at time $t-1$. The $(N+1) \times (N+1)$ covariance matrix at time t conditional upon the filtration $\mathcal{F}_{t-1}$ available at time $t-1$ of the futures holdings with the spot holdings is given by

$$\boldsymbol{\Sigma}_t|\mathcal{F}_{t-1} = \begin{bmatrix} \sigma_{x,t}^2 & \boldsymbol{\Sigma}_{xy,t} \\ \boldsymbol{\Sigma}_{yx,t}^\mathsf{T} & \boldsymbol{\Sigma}_{yy,t} \end{bmatrix} \tag{2.9}$$

which is a partitioned matrix. The block matrix $\boldsymbol{\Sigma}_{yy,t}$ denotes the $N \times N$ conditional covariance matrix of the futures' returns. $\boldsymbol{\Sigma}_{xy,t}$ represents the $1 \times N$ vector of conditional covariances of the spot portfolio returns $\mathbf{X}_{t-1}^\mathsf{T}\gamma_{t-2}$ with every single futures' return $(Y_{1,(t-1)}, ..., Y_{N,(t-1)})$ given the filtration $\mathcal{F}_{t-1}$. Finally, $\sigma_{x,t}^2$ presents the conditional variance of the spot portfolio returns ($\mathbf{X}_{t-1}^\mathsf{T}\gamma_{t-2}$). Since the multivariate distribution of returns might change, $\boldsymbol{\Sigma}_t$ in equation (2.9) is actually an empirical estimation, i.e. $\hat{\boldsymbol{\Sigma}}_t$, with respect to all information available at time $t-1$ of the true covariance for the next period t.

Gagnon et al. (1998) assume that the investor is a mean-variance utility maximiser, which resembles to the following optimisation problem:

$$\underset{\boldsymbol{\delta}}{\operatorname{argmin}} \left[\mathbb{E}[Z_t|\mathcal{F}_{t-1}] - \frac{1}{2}\phi \operatorname{Var}(Z_t|\mathcal{F}_{t-1}) \right] \tag{2.10}$$

where $\phi > 0$ denotes the investor's risk tolerance. By taking the partial derivatives with respect to $\boldsymbol{\delta}$ and setting this equation equal to zero, we obtain the first order condition for a maximum

$$\mathbf{0} = \phi[\boldsymbol{\Sigma}_{yy,t}\boldsymbol{\delta}_t - \boldsymbol{\Sigma}_{xy,t}] - (\mathbb{E}[\boldsymbol{Y}_t|\mathcal{F}_{t-1}] - \boldsymbol{Y}_{t-1}) \tag{2.11}$$

By rearranging the above expression, it is possible to obtain the optimal hedge ratio δ_t conditional upon the information available at time $t-1$, i.e. $\mathcal{F}_{t-1}$

$$\boldsymbol{\delta}_t = \boldsymbol{\Sigma}_{yy,t}^{-1}\boldsymbol{\Sigma}_{xy,t} - \frac{1}{\phi}\boldsymbol{\Sigma}_{yy,t}^{-1}\left(\mathbb{E}[\boldsymbol{Y}_t|\mathcal{F}_{t-1}] - \boldsymbol{Y}_{t-1}\right). \tag{2.12}$$

Notice that $\boldsymbol{\Sigma}_{yy,t}$ has to be non-singular and therefore full-rank. In case some vectors of returns become co-linear, shrinkage methodologies ought to be applied to obtain a robust estimation (see section (3.2)).

From the moment the returns are assumed to be a martingale, the above expression simplifies to

$$\boldsymbol{\delta}_t = \boldsymbol{\Sigma}_{yy,t}^{-1}\boldsymbol{\Sigma}_{xy,t}. \tag{2.13}$$

In other words, the martingale assumption is sufficient for the mean-variance optimal hedge to coincide with the minimum variance hedge in this multi-period setting. Hence, having access to information up and until time $t-1$ is paramount. Notice that this assumption is not far fetched, since it is expected to earn the risk-free rate on money put in the bank account. The theory of interest rate parity (IRP) is essential in foreign exchange markets because it connects interest rates, spot exchange rates, and foreign exchange rates. This theory implies that when the FX market is in a equilibrium, the expected return on domestic assets will equal the exchange rate-adjusted expected return. By the forces of no arbitrage, this essentially means that no premium over the risk-free rate should be expected.

Given foreign exchange market equilibrium, the interest rate parity condition implies that the expected return on domestic assets will equal the exchange rate-adjusted expected return on foreign currency assets. Investors then cannot earn arbitrage profits by borrowing in a country with a lower interest rate, exchanging for foreign currency, an

Due to the importance of the martingale assumption, we refresh the readers' memory. A stochastic process $(Y_{1,t})_{t\in\mathbb{N}}$ on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ is considered a martingale if it satisfies the following conditions:

1. $Y_{1,t}$ is adapted to filtration $(\mathcal{F}_t)_{t\in\mathbb{N}_+}$,
2. $\mathbb{E}[Y_{1,t}|\mathcal{F}_s)] = Y_{1,s})$, for all $s \leq t$ with $s, t \in \mathbb{N}$,
3. $\mathbb{E}[|Y_{1,t}|] < \infty$ for all $t \in \mathbb{N}$.

In other words, it is possible to observe the process $Y_{1,t}$ at every point in time. However, this information will not lead to an increase in the expectation of the process.

2.2.3 Hedging effectiveness

The previous two sections established the general hedging framework for which an estimation has to be calculated. In case the considered stochastic processes are martingale, the mean-variance hedge actually coincides with the minimum variance portfolio. An investor who is a mean-variance utility maximiser will attempt to estimate the empirical covariance matrix and assign the weights to the currencies in his universe that minimise the basket's variance. The question still remains, what is an accurate estimate for the variance-covariance matrix? Besides the classical unbiased sample estimator

$$\hat{\mathbf{\Sigma}} = \frac{1}{T-1}\sum_{t=1}^{T}(\boldsymbol{X}_t - \boldsymbol{\mu}_X)(\boldsymbol{X}_t - \boldsymbol{\mu}_X)^{\mathsf{T}} \tag{2.14}$$

where T is the sample size, $\boldsymbol{\mu}_X$ is the $N \times 1$ mean-vector of the stochastic vector of random variables $\boldsymbol{X}$, section (3.2.0.1), (3.3) and (3.4) present other candidate estimators employed in this thesis.

In reality, hedging is sometimes performed by applying the rule of thumb of selling a predefined fraction of the current FX exposure. Comparing this static hedge to a dynamic hedging strategy requires measure. In addition, we need compare these dynamic hedging strategies, i.e. the different covariance estimators, with each other as well. Traditionally, the hedging effectiveness (HE) index is measured by

$$\text{HE} = \left(\frac{\text{Var}(X) - \text{Var}(Y)}{\text{Var}(X)}\right) \tag{2.15}$$

where the benchmark and hedged portfolios are respectively denoted by X and Y.

The main purpose of this research however is to challenge the iVaR risk measure in the setting of hedging FX exposure. Retail investors in particular wish to have a smooth ride and should not speculate on the FX market since no premiums are to be expected. Using HE as only measure to determine what methodology is superior, is wrong since it is not using variance as objective function. A second measure used in this research is the drawdown reduction index

$$DR = \left(\frac{iVaR(X) - iVaR(Y)}{iVaR(X)} \right). \tag{2.16}$$

Chapter 3

Methodology

3.1 Average drawdown

This section introduces the expected drawdawn which we are optimising for in the iVaR framework. First of all, the notion of an Absolute Drawdown has to be defined. The concept of expected drawdawn is explained thereafter.

3.1.1 Drawdown for a single sample path

The stochastic return vector for an universe of N assets that is discrete in time and continuous in space is denoted by $(\mathbf{X}_t)_{t\in\mathbb{N}} = (X_t^1, ..., X_t^N)_{t\in\mathbb{N}}$ whereas the weights are represented by $(\mathbf{w}_t)_{t\in\mathbb{N}} = (w_t^1, ..., w_t^N)_{t\in\mathbb{N}}$ and the asset prices $(\mathbf{S}_t)_{t\in\mathbb{N}} = (S_t^1, ..., S_t^N)_{t\in\mathbb{N}}$. As mentioned previously, the investment horizon is discrete $\{0, 1, ..., T-1, T\}$ and has $T+1$ subintervals with $T \in \mathbb{N}$ where the uncompounded returns for each single asset are obtained by $X_t^i = S_t^i / S_{t-1}^i - 1$. It easy to compute the in-sample variance-covariance matrix $\mathbf{\Sigma}_T$ for the resulting time series $\mathbf{X}_t$.

At any moment $t \in \{0, 1, ..., T-1, T\}$ the portfolio rate of return is obtained by applying the scalar product $X_t^p = \mathbf{X}_t \cdot \mathbf{w}_t$ for a multi-period model whereas for one-period model $\mathbf{w}$ is independent of t. On the portfolio level, the uncompounded cumulative rate of returns X_t^p for a subinterval t is expressed as

$$Y_t = \begin{cases} 0, & t = 0, \\ \sum_{k=1}^{t} X_t^p, & t = 1, ..., T. \end{cases} \tag{3.1}$$

Chekhlov et al. (2005) then formally defined the absolute drawdown (AD) for a single sample path as

$$\mathrm{AD}(\mathbf{Y}_T) = \boldsymbol{\psi}_T = (\psi_1, ..., \psi_T), \quad \psi_t = \max_{j=\{0,...,t\}} \{Y_j\} - Y_t \tag{3.2}$$

where $\mathbf{Y}_T = (Y_1, ..., Y_T)$ and $\boldsymbol{\psi}_T = (\psi_1, ..., \psi_T)$ are in fact realisations of respectively $(Y_t)_{t\in\mathbb{N}}$ and $(\psi_t)_{t\in\mathbb{N}}$ time series up until time T. Notice that ψ_0 is always equal to zero which is therefore not included in the drawdown time series $\boldsymbol{\psi}$.

Considering the scalar multiplication $\lambda\mathbf{Y}_T = (\lambda Y_1, ..., \lambda Y_T)$ and the scalar addition $a + \mathbf{Y}_T = (a + Y_1, ..., a + Y_T)$ with $\lambda, a \in \mathbb{R}$, Chekhlov et al. (2005) proved that the AD satisfies the properties of a deviation measure, that is

1. Nonnegativity: $\text{AD}(\mathbf{Y}) \geq 0$.
2. Insensitivity to constant shift: $\text{AD}(a + \mathbf{Y}) = \text{AD}(\mathbf{Y})$
3. Positive homogeneity: $\text{AD}(\lambda\mathbf{Y}) = \lambda\text{AD}(\mathbf{Y}),\ \forall\lambda \geq 0$.
4. Convexity: if $Y_t^\lambda = \lambda Y_t^i + (1-\lambda)Y_t^j$ is a linear combination of any two sample paths, i.e. $i, j = 1, ..., N$ and $t = 0, 1, ..., T$ of uncompounded cumulative rates of returns, with $\lambda \in [0, 1]$, then $\text{AD}(Y_t^\lambda) \leq \lambda\text{AD}(Y_t^i) + (1-\lambda)\text{AD}(Y_t^j)$.

The proof of these properties goes as follows: the first three properties are direct consequences from the definition of the AD functional defined in equation (3.2). The last property, on the other hand, is proven by: $\max_{0\leq\tau\leq t}\{\lambda Y_\tau^i + (1-\lambda)Y_\tau^j\} \leq \lambda\max_{0\leq\tau\leq t}\{Y_\tau^i\} + (1-\lambda)\max_{0\leq\tau\leq t}\{Y_\tau^j\}$ with $\lambda \in [0, 1]$.

3.1.2 Minimising the expected drawdown

In order to formally define the expected drawdown (or in sample drawdown) for a sample path, we would, in theory, integrate the probability weighted drawdowns over each state. Our setting is discrete and no probability distribution is assumed beforehand. Therefore, we assume that all observed drawdowns are signals for future drawdowns — sample paths are not noisy — and the historical paths are representative. The expected drawdown is then expressed as

$$\mathbb{E}[\boldsymbol{\psi}_T(\boldsymbol{w})] = \frac{1}{T}\sum_{t=1}^{T}\psi_t(\boldsymbol{w}). \tag{3.3}$$

Remember that $\boldsymbol{\psi}$, defined in equation (3.2), is a non-linear function of the portfolio's weights $\boldsymbol{w}$. Notice that the average of the drawdown process $\boldsymbol{\psi}_T(\boldsymbol{w})$ is in fact the temporal average, i.e. $\mathbb{E}[\boldsymbol{\psi}_T] = \overline{\boldsymbol{\psi}_T}$. Hence, the stochastic process $(\psi_t)_{t\in\mathbb{N}}$ is assumed to be ergodic. However, iVaR is an ex ante optimisation methodology whereas the average drawdown is an ex post risk measure. iVaR is therefore using the ensemble mean by conditioning on paths realised at time t, i.e. $\mathbb{F}_t$.

A few extensions to the temporal average $\frac{1}{T}\sum_{t=1}^{T} a_t\psi_t(\boldsymbol{w})$ are still possible:

1. Exponentially weight a_t the historical drawdowns, i.e. more recent drawdowns are more relevant;

2. Volatility weight the historical drawdowns $a_t = \sigma_t/\overline{\sigma}$;

3. Only include drawdowns when particular conditions are satisfied, e.g. VIX > 30%.

Finding the optimal weights $\mathbf{w}$ that minimise the expected drawdawn, i.e. iVaR, is the solution to the following linear programming (LP) problem

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbb{E}\left[\boldsymbol{\psi}(\mathbf{w})\right] \\ \text{s.t.} \quad & \boldsymbol{\psi}_t = \mathbf{m}_t - \mathbf{w}^{\mathsf{T}}\mathbf{S}_t \\ & \mathbf{m}_t \geq \mathbf{m}_{t-1} \\ & \mathbf{m}_t \geq \mathbf{w}^{\mathsf{T}}\mathbf{S}_t \\ & \mathbf{w}^{\mathsf{T}}\mathbf{1} = 1 \\ & \mathbf{w} \geq \mathbf{0} \end{aligned} \tag{3.4}$$

To ease notation, $\mathbf{S}_t$ represents the realisation of the stochastic vector $(\mathbf{S}_s)_{s\in\mathbb{N}}$ up until time t and has dimensions $N \times t$. Notice that the AD functional could also be written as ψ_t, i.e. $\psi_t = \max\left(\max_{\tau<t}(X^p_\tau) - X^p_t, 0\right)$ which is a non-linear function of the portfolio's path. By introducing the instrumental variable m_t that models the portfolio's monotonic growth, the problem is simplified to an LP.

3.2 Shrinkage of the covariance matrix: Ridge and Lasso

The minimum variance optimisation problem inhibits several drawbacks. First and foremost, we have to estimate the variance-covariance matrix $\hat{\mathbf{\Sigma}}$ from the given time series because the true $\mathbf{\Sigma}$ is unknown. This boils down to the estimation of $N(N+1)/2$ parameters. Therefore, $\mathbf{w}^*$ has to be validated out of sample to mitigate the risk of overfitting since we are using an estimation, i.e. $\hat{\mathbf{\Sigma}}$ in equation 2.2. Next, the estimated variance-covariance matrix $\hat{\mathbf{\Sigma}}$ might become singular when we consider a large number of securities. Equation 2.2 might be unsolvable when $\hat{\mathbf{\Sigma}}$ is not full-rank and thus $\hat{\mathbf{\Sigma}}^{-1}$ doesn't exist. This issue might occur when K is large compared to T. Ledoit and Wolf (2003) state that $\hat{\mathbf{\Sigma}}$ will, most likely, become numerically instable and contain severe estimation errors when the time series' length T is not 10 times larger than the number of assets N.

In statistical modelling it is common to use regularisation techniques to avoid the overfitting of the estimated coefficients. This is achieved by using penalty functions for the weight vector which promotes sparsity among the coefficients. Therefore this paper augments Markowitz's objective function by using penalty functions for the calibration of $\mathbf{w}^*$

$$\underset{\mathbf{w}\in\mathbf{\Theta}}{\arg\min} \left\{ \mathbf{w}^T\mathbf{\Sigma}\mathbf{w} + \lambda \sum_{i=1}^{N} f(w_i) \right\} \tag{3.5}$$

$$\text{subject to} \quad \mathbf{w}^T \mathbf{1} = 1$$

where $f(w_i)$ represents the penalty function to asset i's weight w_i and λ is the regularising constant. The resulting estimation of w^* and σ_p will be both more robust and accurate. By using specific penalty functions, one can also take into account transaction costs induced by trading. In particular the L_2-norm, also known as the RIDGE penalisation, incorporates the impact of order flow imbalances. The L_1-norm, also known as the Least Absolute Shrinkage and Selection Operator (LASSO), on the other hand, considers the bid-ask spread of single trades Caccioli et al. (2016).

3.2.0.1 Regularised portfolio Optimisation: The Elastic Net

In this section, we consider the Elastic Net regularisation method to calibrate a sparse optimal portfolio. This method actually combines the LASSO and the RIDGE method. Therefore, this paper considers de facto three regularising methods. In general, the regulariser essentially restricts the admissable set $\mathbf{w} \in \mathbf{\Theta}$ by penalising large weights. This can be seen in the following equation where the Elastic Net is essentially a weighted average of the L_1-norm and L_2-norm for f in equation 3.5

$$\lambda \sum_{i=1}^{N} f(w_i) = \lambda \sum_{i=1}^{N} \left\{ 0.5 \; (1-\phi) \; w_i^2 + \phi \; |w_i| \right\}, \quad \text{where } \phi \in [0,1]. \tag{3.6}$$

Whenever ϕ is put to one, we obtain the LASSO regularisation method which has several advantages:

1. It controls the total amount of assets that are being shorted. Increasing the values of λ causes the construction of portfolios with less shorting.
2. Sparsity in assets is promoted. In other words, the LASSO penalty performs asset selection because the penalty function is singular at the origin and allows shrinkage to exactly zero.
3. Feature selection mitigates the risk of multicollinear returns.

On the other hand, the disadvantages of LASSO are:

1. If there are more assets than data points, then it selects at most N features;
2. It experiences difficulties in group selection since it will identify only one feature of that group;
3. In general, LASSO is not effective when it is not possible to short sell an asset.
4. The bias generated by LASSO is material when the coefficient in absolute value is large (see Fan and Li (2001).

The Ridge penalisation is obtained when ϕ is equal to zero. The main advantages of this method is that it is able to address the risk of not having a full-rank matrix. In doing so, the RIDGE penalisation will perform grouped selection. On the other hand, it has several disadvantages:

1. It lacks model interpretability.
2. It does not promote sparsity.

3.3 The EWMA covariance matrix

JP Morgan's RiskMetrics used to forecasts covariances and variances using an exponentially weighted moving-average (EWMA) model. This model assigns the largest weights to the most recent observations and lets the weights decrease exponentially over time. It differs significantly with the traditional unbiased estimator of the variance-covariance matrix on the effect of recent events on the volatility compared to events from the distant past. Also, the effect on volatility of a single given observation declines at an exponential rate. To put it differently, the EWMA covariance matrix includes additional information, that is the order of the observations $(X_{1,t}, ..., X_{N,t})$ and $(X_{1,t+1}, ..., X_{N,t+1})$, while this information is neglected in the traditional unbiased estimator. This estimator is furthermore extremely sensitive to the current training set. For instance, if a stressed period would suddenly falls out of the training set, this could have a significant influence. The converse is also true, if the shock is still in the training set while the market regime has clearly changed, the forecasted covariance matrix will remain at an artificially high level.

The EWMA estimate of the covariance matrix is given by

$$\begin{aligned} \hat{\mathbf{\Sigma}}_T &= \frac{(1-\lambda)}{(1-\lambda)^{T-1}} \sum_{t=1}^{T} \lambda^t (\boldsymbol{X}_t - \boldsymbol{\mu}_X)(\boldsymbol{X}_t - \boldsymbol{\mu}_X)^\mathsf{T} \\ &= (1-\lambda)(\boldsymbol{X}_T - \boldsymbol{\mu}_X)(\boldsymbol{X}_T - \boldsymbol{\mu}_X)^\mathsf{T} + \lambda \hat{\mathbf{\Sigma}}_{T-1} \end{aligned} \tag{3.7}$$

where T represents the sample size and the most recent observation and λ, with $0 < \lambda < 1$ denotes the decay factor that determines the relative weight assigned to the observations. RiskMetrics suggests to use $\lambda = 0.94$ for a one-day forecast.

3.4 Forecasting volatility and correlation coefficients

The flexibility feature of dynamically hedging stems from the modelling of the price S_t^i or returns X_t^i time series. This section will therefore first introduce the most important elements of time series analyses. Next, the (generalised) autoregressive conditional heteroskedasticity is formally defined to model the time dependent stochastic volatility parameter.

3.4.1 The building blocks for time series analysis

As mentioned previously, we observe a discrete-time stochastic process $(X_t)_{t\in\mathbb{N}}$ which is defined on the filtered probability space $(\Omega, \mathcal{F}, P)$. The definitions listed below, given by **?**, try to formally capture the idea that the stochastic properties (e.g. distribution, moments etc.) of a time series are not influenced by the moment one observes it.

Definition 3.4.1 (strong stationarity)**.** *The time series $(X_t)_{t\in\mathbb{N}}$ is strictly stationary if*

$$(X_t, X_{t+1}, ..., X_{t+k-1}, X_{t+k}) \stackrel{\mathrm{d}}{=} (X_s, X_{s+1}, ..., X_{s+k-1}, X_{s+k})$$

for all t, s and $k \in \mathbb{N}$.

Definition 3.4.2 (weak stationarity or covariance stationarity)**.** *The time series $(X_t)_{t\in\mathbb{N}}$ is weak stationary if the first two moments exist and satisfy*

$$\begin{aligned} &\mu(t) = \mu, \quad t \in \mathbb{N}, \\ &\gamma(t,s) = \gamma(t+k, s+k); \quad t, s, k \in \mathbb{N}, \end{aligned}$$

where

$$\begin{aligned} &\mu(t) = \mathbb{E}[X_t], \quad t \in \mathbb{N}, \\ &Cov(X_t, X_s) = \gamma(t,s) = \mathbb{E}[(X_t - \mu(t))(X_s - \mu(s))], \quad t, s \in \mathbb{N}. \end{aligned}$$

It is clear that the auto-covariance function is symmetric $\gamma(s,t) = \gamma(t,s)$ for all $s, t \in \mathbb{N}$ and that $\gamma(t,t) = \mathrm{Var}[X_t]$. Moreover, the notation for covariance - stationary time series simplifies: $\gamma(h) = \gamma(h, 0)\, \forall h \in \mathbb{N}$ and $\gamma(0) = \mathrm{Var}(X_t)\, \forall t$. A strictly stationary time series with $\mathrm{Var}[X_t] < \infty$ is also covariance stationary. Note that some (G)ARCH stochastic processes have infinite variance while they are strictly stationary still, but are not covariance stationary. To put the previous two definitions differently, a strictly/weakly stationary time series must not have any systematic changes in mean, variance or covariances between equally spaced observations.

Definition 3.4.3 (autocorrelation function)**.** *The autocorrelation function of $(X_t)_{t\in\mathbb{N}}$ for lag h is*

$$\rho(h) = \rho(X_h, X_0) = \gamma(h)/\gamma(0), \quad \forall h \in \mathbb{N}.$$

Definition 3.4.4 (white noise)**.** *$(X_t)_{t\in\mathbb{N}}$ is considered to be a white noise process if it is both covariance stationary and the correlation function satisfies*

$$\rho(h) = \begin{cases} 1, & h = 0, \\ 0, & h \neq 0. \end{cases}$$

A good example of a white noise process is the strict white noise process.

Definition 3.4.5 (strict white noise)**.** *$(X_t)_{t\in\mathbb{N}}$ is a strict white noise process if it is a sequence of independent identically distributed random variables with finite variances.*

In order to introduce martingale difference, we discuss the sigma algebra $\mathcal{F}_t = \sigma(\{X_s : s \leq t\})$ which contains all information of the stochastic process $(X_t)_{t\in\mathbb{N}}$ up until time t. To put it differently, it models the accrual of information over time. This implies that the indexed sequence of sigma algebras $(\mathcal{F}_t)_{t\in\mathbb{N}}$ is the natural filtration since $\sigma(\{X_s : s \leq t\}) \subseteq \sigma(\{X_s : s \leq t+1\})$.

Definition 3.4.6 (martingale difference)**.** *The time series $(X_t)_{t\in\mathbb{N}}$ is a martingale difference series if it is adapted to filtration $(\mathcal{F}_t)_{t\in\mathbb{N}}$, integrable $\mathbb{E}[|X_t|] < \infty$ and centered $\mathbb{E}[X_t|\mathcal{F}_{t-1}] = 0$, $\forall t \in \mathbb{N}$.*

By applying the tower rule (the law of total expectation), we notice that the unconditional expectation with respect to the stochastic process $(X_t)_{t\in\mathbb{N}}$ is zero as well: $\mathbb{E}[X_t] = \mathbb{E}[\mathbb{E}[X_t|\mathcal{F}_{t-1}]] = 0$, $\forall t \in \mathbb{N}$. A fortiori, if the process has a finite second raw moment, i.e. $\mathbb{E}[X_t^2] < \infty$, the covariance is also equal to zero

$$\gamma(t,s) = \mathbb{E}[X_t X_s] = \begin{cases} \mathbb{E}[\mathbb{E}[X_t X_s|\mathcal{F}_{s-1}]] = \mathbb{E}[X_t \mathbb{E}[X_s|\mathcal{F}_{s-1}]] = 0, & t < s, \\ \mathbb{E}[\mathbb{E}[X_t X_s|\mathcal{F}_{t-1}]] = \mathbb{E}[X_s \mathbb{E}[X_t|\mathcal{F}_{t-1}]] = 0, & s < t. \end{cases}$$

This section can be concluded by a observation that is using all the stated definitions. Assuming that we have a finite variance martingale-difference process which is therefore integrable, adapted to filtration, centered and has zero covariance, is a white noise process if the variance is constant throughout time t.

3.4.2 Autoregressive Conditional Heteroskedasticity

The basic univariate GARCH is introduced first in order to generalise thereafter to a multivariate version. **?** define that $(X_t)_{t\in\mathbb{N}}$ is a GARCH(p,q) if it is strictly stationary and it satisfies the equations

$$\begin{aligned} X_t &= \sigma_t Z_t, \quad \forall t \in \mathbb{Z} \\ \sigma_t^2 &= c + \sum_{i=1}^{p} \alpha_i (X_{t-i})^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \quad \forall t \in \mathbb{Z}, \end{aligned} \tag{3.8}$$

where $c > 0$, $\alpha_i \geq 0$, $i = 1, ..., p$ and $\beta_j \geq 0$, $j = 1, ..., q$. It is called a generalised autoregressive process since it not only depends on the past volatilities but also on the history of the squared process (X_t).

Lee and Yoder (2007) define a multivariate GARCH(1,1) model for two discrete-time stochastic processes which we will generalise for our universe of N assets, that is $(X_{i,t})_{t\in\mathbb{N}}$ where $i = 1, ..., N$. In matrix notation, this resembles to $(\boldsymbol{X}_t)_{t\in\mathbb{N}}$. First, we assume that the correlation coefficients between these asset are constant throughout time. Among others, Engle (2002) employed also a constant correlation coefficient (CCC) GARCH(1,1) model, developed by Bollerslev (1990), for the purpose of hedging a portfolio with exposure to various foreign currencies. This model enables the hedge ratio to be time-varying and includes the long-term relationship between different spot currency prices. The multivariate GARCH(1,1) is formally expressed as

$$\boldsymbol{X}_t = \boldsymbol{\mu}_t + \boldsymbol{a}_t \tag{3.9}$$

where $\boldsymbol{\alpha}_t$ is the expected return vector at time t and

$$\boldsymbol{a}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_t). \tag{3.10}$$

Since linear transformations of normal random variables are also normally distributed, it is possible to express $\boldsymbol{a}_t$ as

$$\boldsymbol{a}_t = \boldsymbol{\Sigma}_t^{1/2} \boldsymbol{z}_t \tag{3.11}$$

where $\boldsymbol{z}_t$ is a $N \times 1$ vector of identically independent distributed standard normal variables, i.e. $\mathbb{E}[\boldsymbol{z}_t] = 0$ and $\mathbb{E}[\boldsymbol{z}_t \boldsymbol{z}_t^\intercal] = \boldsymbol{I}_N$. The Cholesky decomposition of $\boldsymbol{\Sigma}_t$ is expressed by $\boldsymbol{\Sigma}_t^{1/2}$.

$\boldsymbol{\Sigma}_t$ represents the conditional $N \times N$ covariance matrix at time t. For instance, consider two assets $(X_{1,t})_{t\in\mathbb{N}}$ and $(X_{2,t})_{t\in\mathbb{N}}$ where this conditional covariance matrix can be decomposed as

$$\begin{aligned}
\boldsymbol{\Sigma}_t &= \begin{bmatrix} \sigma_{11,t}^2 & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t}^2 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_{11,t} & 0 \\ 0 & \sigma_{22,t} \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} \sigma_{11,t} & 0 \\ 0 & \sigma_{22,t} \end{bmatrix} \\
&= \mathbf{D}_t \mathbf{R} \mathbf{D}_t
\end{aligned} \tag{3.12}$$

As seen before, the GARCH(1,1) model is in fact nothing more than an ARMA(1,1) process of the squared residuals, that is

$$\begin{aligned}
\sigma_{1,t}^2 &= c_1 + \alpha_1 \epsilon_{1,t-1}^2 + \beta_1 \sigma_{1,t-1}^2 \\
\sigma_{2,t}^2 &= c_2 + \alpha_2 \epsilon_{2,t-1}^2 + \beta_2 \sigma_{2,t-1}^2.
\end{aligned} \tag{3.13}$$

In the general setting, the N-dimensional white noise with conditional variance - covariance matrix $\boldsymbol{\Sigma}_t$ is modeled by $\boldsymbol{a}_t = (a_{1,t}, ..., a_{N,t})^\intercal$ and follows a centered N-dimensional normal distribution. The pairwise Pearson correlation coefficients $\rho_{i,j}$ for $i, j = 1, ..., N$ are assumed to be constant.

The Dynamic Correlation Coefficient GARCH (DCC GARCH) introduced by **?** and Engle (2002) relaxes the assumption of time independent correlation coefficients among random variables.

$$\mathbf{R}_t = \mathbf{J}_t^{-1} \mathbf{Q}_t \mathbf{J}_t^{-1} \tag{3.14}$$

where

$$\mathbf{J}_t = \begin{bmatrix} \sqrt{q_{11,t}} & 0 & \dots & 0 \\ 0 & \sqrt{q_{11,t}} & \dots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sqrt{q_{NN,t}}. \end{bmatrix} \tag{3.15}$$

Remember that $\boldsymbol{D}_t$ represents the diagonal matrix containing the conditional standard deviations of every asset at time t. By performing a linear transformation on $\boldsymbol{a}_t$, we obtain standardised $N \times 1$ vector of standardised residuals where the conditional covariance matrix equals the conditional correlation matrix,

$$\boldsymbol{\epsilon}_t = \boldsymbol{D}_t^{-1}\boldsymbol{a}_t \sim \mathcal{N}(0, \boldsymbol{R}_t). \tag{3.16}$$

Finally, $\mathbf{Q}_t$ which is positive definite, is defined as an autoregressive moving average process

$$\mathbf{Q}_t = (1 - \theta_1 - \theta_2)\bar{\mathbf{Q}} + \theta_1 \boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}^{\mathsf{T}} + \theta_2 \mathbf{Q}_{t-1}. \tag{3.17}$$

The parameters θ_1 and θ_2 are scalars which have to satisfy

$$\theta_1 \geq 0, \theta_2 \geq 0 \quad \text{and} \quad \theta_1 + \theta_2 < 1. \tag{3.18}$$

The unconditional covariance matrix is denoted by $\bar{\mathbf{Q}} = \mathbb{E}[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^{\mathsf{T}}]$ and is estimated by

$$\bar{\mathbf{Q}} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^{\mathsf{T}}. \tag{3.19}$$

Notice that $\boldsymbol{J}_t$ rescales the elements of $\boldsymbol{Q}_t$ such that $\left|\dfrac{\sigma_{ij,t}^2}{\sigma_{ii,t}\sigma_{jj,t}}\right| = \rho_{ij,t} \leq 1$.

3.4.3 Optimisation for CCC and DCC GARCH(1,1)

Both CCC and DCC GARCH models implicitly assume a distribution for the residuals. In the literature, distributions as the multivariate Gaussian, Student's t and skew Student's t distributions are used. Nevertheless, this research will only consider multivariate Gaussian distribution.

3.4.3.1 Multivariate Normal distribution

This research assumes that the vector of independent standardised residuals $[\boldsymbol{z}_0, ..., \boldsymbol{z}_T]$ are multivariate Gaussian distributed, that is the joint distribution

$$f(\boldsymbol{z}_t) = \prod_{t=1}^{T} \frac{1}{(2\pi)^{N/2}} \exp\left\{-\frac{1}{2}\boldsymbol{z}_t^\mathsf{T}\boldsymbol{z}_t\right\} \tag{3.20}$$

where $\mathbb{E}[\boldsymbol{\epsilon}_t] = 0$ and $\mathbb{E}[\boldsymbol{z}_t^\mathsf{T}\boldsymbol{z}_t] = \boldsymbol{I}_N$ with $t = 1, ..., T$ representing the observations in the sample period.

For normal distributed random variables we know that a linear transformation, i.e. $\boldsymbol{a}_t = \boldsymbol{\Sigma}_t^{1/2}\boldsymbol{z}_t$ is again normally distributed and the likelihood function resembles to

$$L(\boldsymbol{\Theta}) = \prod_{t=1}^{T} \frac{1}{(2\pi)^{N/2}|\boldsymbol{\Sigma}_t|} \exp\left\{-\frac{1}{2}\boldsymbol{a}_t^\mathsf{T}\boldsymbol{\Sigma}_t^{-1}\boldsymbol{a}_t\right\}. \tag{3.21}$$

The parameter set $\boldsymbol{\Theta}$ contains the model parameters for the CCC and DCC GARCH(1,1) model. In particular, the parameter set contains both the univariate GARCH parameters $(c_i, \alpha_{1,i}, \beta_i)$ for $i = 1, ..., N$ and the correlation parameters (θ_1, θ_2). The goal is to find the parameter set that maximises the above likelihood function. By substituting $\boldsymbol{\Sigma}_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t$ and taking the natural logarithm, we obtain

$$\begin{aligned}\ln(L(\boldsymbol{\Theta})) &= -\frac{1}{2}\sum_{t=1}^{T}\Big(N\ln(2\pi) + \ln(|\mathbf{D}_t\mathbf{R}_t\mathbf{D}_t|) + \boldsymbol{a}_t^\mathsf{T}t(\mathbf{D}_t\mathbf{R}_t\mathbf{D}_t)^{-1}\boldsymbol{a}_t\Big)\\ &= -\frac{1}{2}\sum_{t=1}^{T}\Big(N\ln(2\pi) + 2\ln(|\mathbf{D}_t|) + \ln(|\mathbf{R}_t|) + \boldsymbol{a}_t^\mathsf{T}\mathbf{D}_t^{-1}\mathbf{R}_t^{-1}\mathbf{D}_t^{-1}\boldsymbol{a}_t\Big).\end{aligned} \tag{3.22}$$

Notice that in the above likelihood function not only the variance coefficients are time dependent, but also the correlation coefficients. This significantly complicates the optimisation problem which is why it is split into two stages. First, the parameter set that corresponds to the univariate GARCH(1,1) models are estimated by substituting $\mathbf{R}_t$ by $\mathbf{I}_N$. The parameters corresponding to the correlation coefficients (θ_1, θ_2) are then estimated in the second stage. Notice that for the CCC GARCH(1,1) we actually only need to perform the first stage and then calculate the Pearson correlation matrix for all the assets N.

Step one: univariate GARCH(1,1)

As mentioned previously, the $\mathbf{R}_t$ is substituted by $\mathbf{I}_N$ in the log-likelihood function

$$\begin{aligned} \ln(L_1(\boldsymbol{\Theta})) &= -\frac{1}{2}\sum_{t=1}^{T}\Big(N\ln(2\pi) + 2\ln(|\mathbf{D}_t|) + \ln(|\mathbf{I}_N|) + \boldsymbol{a}_t^{\mathsf{T}}\mathbf{D}_t^{-1}\mathbf{I}_N\mathbf{D}_t^{-1}\boldsymbol{a}_t\Big)\,. \\ &= -\frac{1}{2}\sum_{t=1}^{T}\left(N\ln(2\pi) + \sum_{n=1}^{N}\left\{\ln(\sigma_{i,t}) + \frac{a_{i,t}^2}{\sigma_{i,t}}\right\}\right)\,. \end{aligned} \tag{3.23}$$

Appendix A.1 elaborates on the univariate maximum likelihood estimation of the GARCH(1,1) model. Thus, the last equation is merely the sum of every individual log-likelihood function of a GARCH(1,1) for N assets. This can be seen as a quasi-likelihood estimation for a multivariate GARCH(1,1) that are not correlated. Once the set of parameters $(c_i, \alpha_{1,i}, \beta_i)$ is determined for $i = 1, ..., N$, it is possible to calculate $\bar{\mathbf{Q}} = \mathbb{E}[\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t^{\mathsf{T}}]$ and $\boldsymbol{\epsilon}_t = \mathbf{D}_t^{-1/2}\boldsymbol{a}_t$.

Step two: correlation coefficients

The set (θ_1, θ_2) has still to be determined by the following quasi-likelihood function

$$\begin{aligned} \ln(L_2(\boldsymbol{\Theta})) &= -\frac{1}{2}\sum_{t=1}^{T}\Big(N\ln(2\pi) + 2\ln(|\mathbf{D}_t|) + \ln(|\mathbf{R}_t|) + \boldsymbol{a}_t^{\mathsf{T}}\mathbf{D}_t^{-1}\mathbf{R}_t^{-1}\mathbf{D}_t^{-1}\boldsymbol{a}_t\Big)\,. \\ &= -\frac{1}{2}\sum_{t=1}^{T}\Big(N\ln(2\pi) + 2\ln(|\mathbf{D}_t|) + \ln(|\mathbf{R}_t|) + \boldsymbol{\epsilon}_t^{\mathsf{T}}\mathbf{R}_t^{-1}\boldsymbol{\epsilon}_t\Big)\,. \\ &= -\frac{1}{2}\sum_{t=1}^{T}\Big(\text{constant} + \ln(|\mathbf{R}_t|) + \boldsymbol{\epsilon}_t^{\mathsf{T}}\mathbf{R}_t^{-1}\boldsymbol{\epsilon}_t\Big)\,. \end{aligned} \tag{3.24}$$

First, notice that $\mathbf{D}_t$ is constant because it was estimated in the first stage. **?** showed that this pseudo-likelihood method has asymptotically normal estimators. In addition, Jondeau and Rockinger (2005) proved that this two stage approach is similar to an one-step full maximum likelihood estimation.

3.4.3.2 Potential pitfalls

Since this is a multivariate optimisation problem, the initial starting values are extremely important in order to avoid local optima. It is therefore of uttermost importance to use a grid of initial values to be sure that the final set of parameter values is a global optimum. A second potential pitfall is the existence of outliers in the sample set. Gradient optimisation algorithms might run into trouble when these observations are not removed, i.e. it hits a boundary. The Sequential Least Squares numerical optimisation algorithm was used in accordance with the SciPy package in python: `scipy.optimize.minimize(method='SLSQP')`.

3.5 Hansen's model confidence set

The goal is to determine the set of optimal model(s) V^* which outperforms the whole set of considered models, i.e. V^0. Hansen's Model Confidence Set (MSC) (Hansen et al. (2011)) calculates these p-values by choosing the best-in-class models that aren't significantly different from each other based on a predetermined metric.

The Hansen's MSC is a series of significance tests where the null hypothesis is

$$H_{0,V} : \mathbb{E}[d_{ij,t}] = 0, \quad \forall i, j \in V^0. \tag{3.25}$$

The loss function $d_{ij,t} = L_{i,t}L_{j,t}$ is calculated for two models i and j in the complete set of models $i, j \in V^0$ at a specific observation date t in the backtest. The alternative models are then ranked according to the expected loss. For instance, model i is to be preferred if $\mathbb{E}[d_{ij,t}] < 0$. The total set of optimal models is then obtained by

$$V^* = \{i \in V^0 \,|\, \mathbb{E}[d_{ij,t}] \leq 0, \, \forall j \in V^0\}. \tag{3.26}$$

Finally, the metric to compare these models over for this research are: returns, volatility, Sharpe Calmar and pain ratio, maximum and average drawdown. For each of these criteria, the value is calculated for each of the portfolios at each month over the duration of the backtest, resulting in a list of values that can then be compared over time using Hansen's statistical test. The return statistic is calculated as the simple monthly percentage change in portfolio value. The maximum drawdown is calculated as the maximum of the percentage drawdown from the high watermark at that time. Equally, the average drawdown is calculated as the mean over time of the percentage drawdown from the high watermark.The Calmar ratio is calculated as the monthly return divided by the monthly maximum drawdown. The pain ratio is calculated as the monthly return divided by the monthly average drawdown.

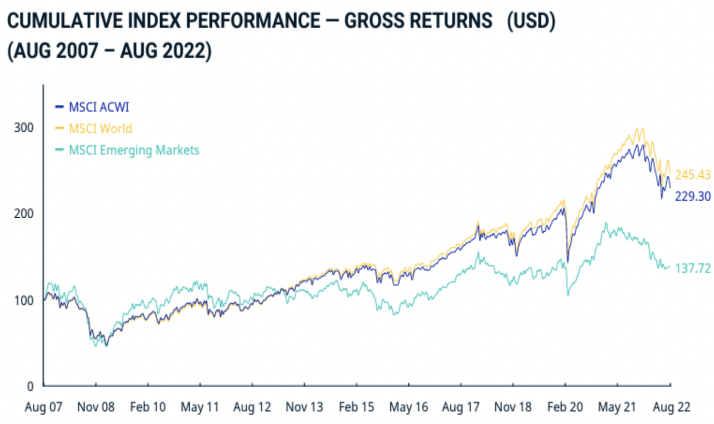
Chapter 4

Numerical experiment

4.1 Case study: MSCI ACWI

This research investigates different FX hedging methods on the Morgan Stanley Capital International (MSCI) All Country World Index (ACWI), that is MSCI's flagship global equity index. Its investment universe ranges from large - to mid-cap stocks across 23 developed and 24 emerging markets. For instance, on May 2022, its exposure was almost as large as 3000 constituents that covered almost 85 % of the free float-adjusted market capitalisations in more than ten sectors[1]. Figure (4.1) and (4.1) show both the time series of MSCI ACWI and the different sectors it is composed of.

[1] A full description of MSCI ACWI is available at: https://www.msci.com/our-solutions/indexes/acwi



ANNUAL PERFORMANCE (%)

Year	MSCI ACWI	MSCI World	MSCI Emerging Markets
2021	19.04	22.35	-2.22
2020	16.82	16.50	18.69
2019	27.30	28.40	18.88
2018	-8.93	-8.20	-14.24
2017	24.62	23.07	37.75
2016	8.48	8.15	11.60
2015	-1.84	-0.32	-14.60
2014	4.71	5.50	-1.82
2013	23.44	27.37	-2.27
2012	16.80	16.54	18.63
2011	-6.86	-5.02	-18.17
2010	13.21	12.34	19.20
2009	35.41	30.79	79.02
2008	-41.85	-40.33	-53.18

Figure 4.1: MSCI ACWI time series until August 2022.

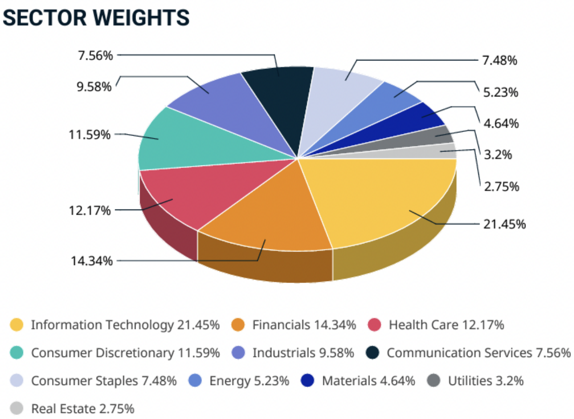


Figure 4.2: Sector weights for the MSCI ACWI index on August 2022.

4.2 From raw data to results

In order to perform a successful backtest, a long time series is predominant in order to thoroughly evaluate the chosen optimisation methodologies out of sample. First line of business is the extraction of currency exposure data for different time stamps. This is accomplished by first using free float-adjusted market capitalisations for every single instrument and subsequently by grouping per home currency. MSCI publishes ISIN data for its ACWI index back until January 2004 which therefore determines the start of this research's backtesting period. As one can expect, these currency exposures vary over time because market capitalisations will. This is displayed in figure (4.2).

From a cost perspective, it is not interesting to hedge every FX exposure. Currency exposures that could be considered immaterial are most likely correlated with other currencies where the market might just be not as liquid. This wide bid-ask spread will most surely offset all the potential diversification benefits. The universe of currencies is therefore restricted to : `[USD, EUR, JPY, GBP, HKD, CAD, CHF, AUD, TWD, INR, CNY, SAR, BRL, KRW, ZAR, RUB, SEK, NOK]`. A table of the entire list can be found in appendix (C.1).

The optimisation problem has to be clearly defined still. In particular, the bounds set to each single currency in the universe as well as the maximum exposure that we could hold during the invested period are crucial. In case the bounds are too strict, we risk of having either an infeasible solution or to not have an optimisation problem at all. Consider the case where an Euro denominated investor would like to hedge his FX exposure inhibited in the invested basket. Say that he is not particularly risk averse and would like to keep 15% EUR at all time. As figure (4.2) already showed, Euro's contribution to MSCI ACWI is ranging from 10% to 20% throughout time which implies that the problem becomes infeasible when $w_{t,EUR} > 15\%$ for $t = 1, ..., T$. One might wonder, if we just sell some Euros for foreign currencies the issue should be remediated and the initial optmisation problem becomes feasible again. This might be the case for equities or other financial instruments, but it is hard to explain to investors that our optimiser would actually buy more foreign currencies in order to reach our hedging goal which is in essence to reduce risk FX risk. Therefore strict upper bounds are imposed on each individual currency that it is only possible to assign a weight as large as 100% of MSCI's magnitude in that particular currency. The weights also have to sum to one, i.e. $\boldsymbol{w}^\intercal \mathbf{1} = 1$. Lastly, it is not possible to sell more than the current exposure: $w_i \geq 0$ for $i = 1, ..., N$. Coming back to the illustrative problem where infeasibility might occur, it is possible to just hold 20% EUR at all times. This would cause that at some particular time stamps we don't have an optimisation problem at all: $w_{t,EUR} \approx 20\%$ but always less than 20%. This potential pitfall is avoided by altering the definition of hedging. Instead of keeping a fixed weight in Euro throughout time, a fixed fraction of the total foreign currency exposure is sold at predetermined time stamps. For instance, if the total FX weight amounts to 80% at one point in time, half of this exposure is sold. The total share of Euros held by this investor amounts to 60%. This is exactly the 50% static hedging benchmark this research considers.

After handling the raw ISIN data, determining a fixed universe of currencies and defining the optimisation problem, it is possible to construct the backstest for different optimisation methods and benchmark it against the 50% static hedged basket. The backtesting period is ranging from January 2004 until January 2022 which includes particularly interesting events:

1. The financial crises of 2008.
2. The sovereign debt crisis.
3. The unpegging of the Swiss Franc to the Euro.
4. The unpegging of the Chinese Yuan to the USD.
5. The covid crisis.

Next to the range of the backtesting period, it is important to have an appropriate calibration period. On the one hand, it cannot be too short since there is a significant risk of overfitting the data or the possibility that some covariance methods might become not full rank anymore. By using an extremely long calibration period, on the other hand, we might not be able to capture changing trends, due to the arrival of new information, in the data set. Therefore this research is using three years of data as training set and one month as test set. After one month, the calibration is run again and rebalancing takes

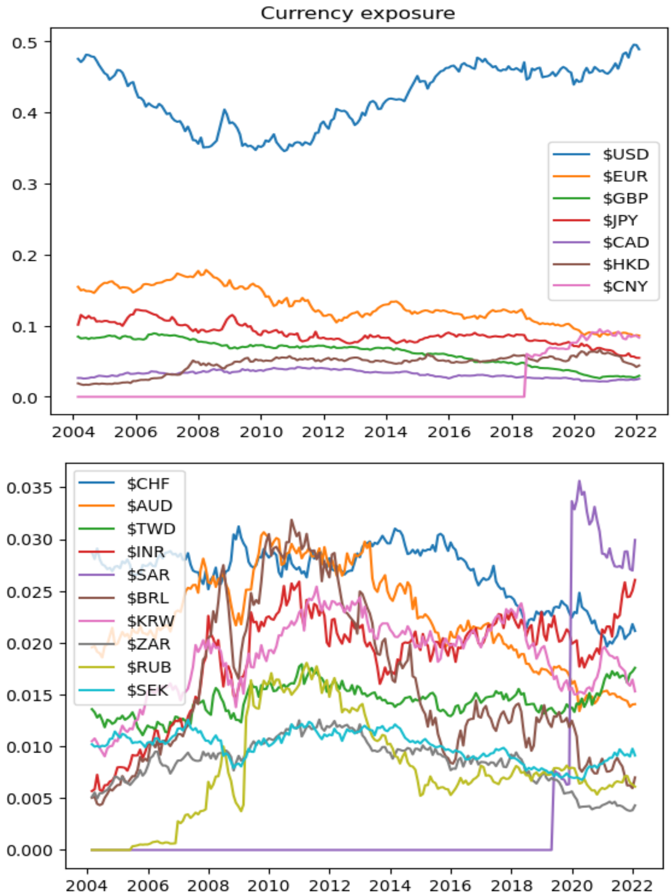


Figure 4.3: It is clear that currencies exposures vary significantly throughout time. For instance, the Chinese Yuan and the Saudi Arabian Riyals were only relevant from 2018 and 2019 respectively. The British Pound Sterling and EURO, on the other, diminished throughout the years. The United States Dollar exposure in the MSCI ACWI varies from 40% to 50% and is by far the largest risk driver in the basket.

place both in the benchmark as in the optimal portfolio. Note that during the backtesting loop, the exposures of every currency at the start of the one month test set, are given as upper bounds to the MOSEK[2] package. By dynamically keeping track of these exposures, it is possible to realistically compare the performance of the optimal portfolio with the benchmark. In reality, hedging activities in retail portfolios will only happen after a fixed interval of time, e.g. one month, and not daily.

4.3 Descriptive statistics

This section summarises the descriptive statistics of the daily return time series denominated in Euro since this is our base case (see next section). These findings can be generalised to the other cases as well (i.e. USD, GBP). Table (4.1) summarises the mean, standard deviation, skewness, kurtosis, the Augemented Dicky Fuller (ADF) test and fi-

[2]A full description of the MOSEK package is available at: https://www.mosek.com/documentation/

nally the Jarque-Bera Lagrange multiplier statistic for the daily returns of all currencies in the considered universe.

Standard in time series analysis is to start with an examination of unit roots. The ADF is employed to test for all currency returns in each market that under the null hypothesis of a unit root against the alternative hypothesis of not containing an unit root. This test rejects the null-hypotheses at 1% significance level, meaning that all return series could potentially be stationary since they do not contain an unit root. Furthermore, we can conclude that the mean return for all currency pairs with EUR are statistically not different from 0 because no trend is observed anymore. In section (2.2.2) it was already assumed that there is no premium for returns in the FX market. The corresponding variance of returns, on the contrary, is much higher. Figure (4.4) presents four graphs of daily returns for EURUSD, EURGBP, EURCHF and EURKRW. These indicate volatility clustering, or periods of high volatility followed by periods of relative tranquility. In particular, all graphs inhibit high volatility periods in 2008, 2011, 2016 and 2020. This also shows that these time series are at least highly correlated in stressed periods. This research tries to account for changing volatility regimes by estimating a currency pair's standard deviation via an ARMA(1,1)-like process (autoregressive moving average model for the error's variance), see section (3.4).

These currencies' spot returns series have also a high kurtosis, which indicates the presence of fat tails. Especially the currency pairs: EURCHF, EURAUD, EURGBP, EURJPY, EURKRW, EURRUB and EURZAR. On the other hand, EURUSD seems to be thin-tailed and by incorporating the skewness, also symmetrical. This is an exception to most currency pairs since they all are either large negative or positive skewness statistics. This indicates a longer left tail (extreme losses) for a negative coefficient and longer right tail for a positive coefficient (extreme gains). The EURCHF has for instance an extreme positive outlier. This was the result of a change in the Swiss' central bank policy by abolishing the peg to the Euro. In appendix (B.1) the correlation and co-drawdown matrix as well as the Principal Component Analysis are displayed. Notice that the first four eigenvectors of the correlation matrix already explains 71% of the variance, that is of the 16 currency pairs. As we know that the percentage explained variance of each eigenvector can be calculated by $\frac{\lambda_i}{\sum_i^N \lambda_i}$, the first four eigenvectors explain: $[45\%, 15\%, 6\%, 5.7\%]$. In other words, it seems that these currency pairs are linearly dependent.

	Mean	SD	Skewness	Kurtosis	ADF Test	Jarque–Bera
AUD	0.000026	0.006683	-0.421104	12.931220	-17.544961	32648.37
BRL	-0.000054	0.010217	-0.028368	4.446663	-52.341691	3843.40
CAD	0.000050	0.005729	-0.065047	2.389531	-49.601549	1112.43
CHF	0.000100	0.004943	13.625825	628.554538	-15.221178	76971210.17
CNY	0.000095	0.005717	-0.019480	2.357166	-71.356732	1079.57
GBP	-0.000035	0.005070	-0.426934	6.261091	-49.337322	7761.89
HKD	0.000038	0.005684	-0.011969	2.357484	-68.765878	1079.68
INR	-0.000064	0.006351	-0.034825	2.672162	-30.097996	1388.13
JPY	0.000037	0.007134	0.407986	10.965139	-13.810971	23504.79
KRW	0.000047	0.007612	0.384157	25.389884	-12.690677	125459.33
RUB	-0.000149	0.009089	0.219577	33.747277	-10.472293	221485.55
SAR	0.000039	0.005727	-0.024494	2.482098	-68.727603	1197.25
SEK	-0.000017	0.004337	-0.176943	4.671136	-13.075981	4265.05
TWD	0.000078	0.005735	-0.019409	2.314423	-42.451837	1040.76
USD	0.000039	0.005710	0.000571	2.325973	-68.508853	1050.89
ZAR	-0.000111	0.009645	-0.883822	11.535827	-28.832720	26480.39

Table 4.1:

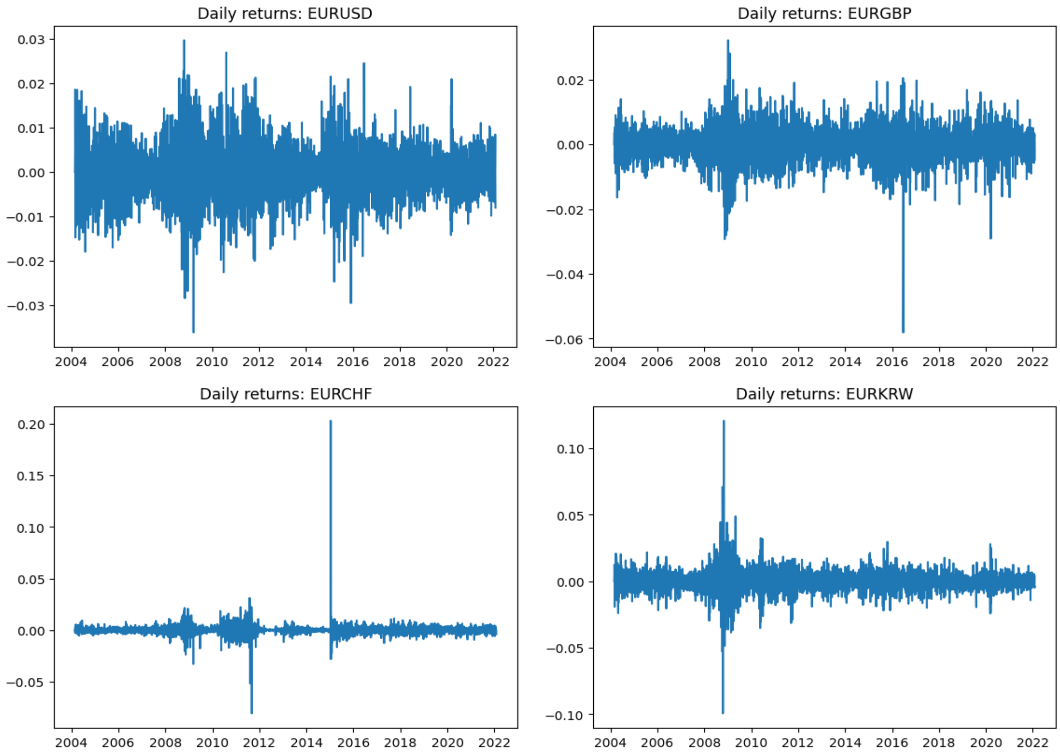


Figure 4.4: test

4.4 Evaluation criteria

In order to thoroughly evaluate the overall performance of different optimisation methodologies with the iVaR framework, it is necessary to adapt a statistical framework. As already formally introduced in section (3.5), Hansen's Model confidence set is an approach to determine which models under- or outperform relative to others, giving the objective: average drawdown, volatility, Sharpe ratio, calmar and pain ratio. Since the obtained results will depend on the overall setting, this research considers:

1. The base case where the static Euro denominated 50% hedged benchmark portfolio is compared against the iVaR and the minimum variance, computed using the historical covariance matrix.

2. The home currency might be relevant which is why the analysis of the base case from above is also performed on USD and GBP denominated portfolios. The former plays an incremental role in the basket since it almost 50% of the MSCI ACWI stocks are denominated in USD. On the other hand, GBP's fraction is much lower than USD and EUR, especially near the end of the backtest period.

3. As in many quantitative methodologies applied to finance, transaction costs are assumed to be negligible or even not considered at all. In the event they are taken into account, it in general offsets the whole benefit. This research therefore considers transaction costs in the rebalancing after the one month test set. After the calibration, the optimal weights are compared to the previous ones, that is of one month earlier, and only a trade occurs if it offsets the imposed transaction cost.

4. We already are aware of the significant limitations the historical covariance's estimation, hence these results are challenged by the constant and dynamic correlation coefficient approach but also the shrunken covariance estimation (L1 and L2) and lastly the exponentially weighted moving average covariance matrix.

4.5 Results

The results obtained in this research indicate that the methodology for hedging foreign currency risk is highly dependent on the problem setting from an optimisation point of view. In case strict bounds are imposed on the individual weights, the performance of an iVaR portfolio is comparable to a minimum variance portfolio. Both portfolios perform better than the static 50% hedged benchmark portfolio in terms of hedging effectiveness and drawdown reduction, but do not differ significantly from each other. This is not due to ill-defined methodologies, but rather due to the specifics of the case study, especially in the base case. This proven by having a de facto unconstrained optimisation problem. When using a different base currency, iVaR slightly performs better than the minimum variance portfolio, but more importantly, still performs better than the static hedge. Transaction costs seem not to play an important role for the iVaR portfolio in the performance with respect to the cumulative drawdowns. Losses, however do occur more often, but these are still materially different from the static hedge. Lastly, different covariance estimators have been tested in order to challenge the unbiased estimator of the covariance matrix since the minimum variance optimiser is extensively used throughout this thesis. This research finds that in this strict setting, the different estimators are not statistically different from each other according to Hansen's model confidence set.

4.5.1 Base case: static hedge, iVaR and minimum variance

The obtained backtesting results for the static 50% hedged benchmark, the iVaR and the historical minimum variance estimate are presented in figure (4.5). A linear scale was used for the y-axis, which is also the case for the figures that will follow. Also the portfolio values are standardised such that they all start from 1.00. It is clear that both the iVaR and the minimum variance optimisation methods are able to make timeseries more monotone, for example in the stressed period of 2008. This is also the case during the European sovereign debt crisis in 2010-2011. The iVaR on the other hand, managed to get out faster of the 2013-2015 valley. As already mentioned in section (4.3), the CHF experienced a sudden spike on 15 January 2O15 which explains the sharp increase in all portfolios.

The histogram in figure (4.5) displays the cumulative drawdowns, that is the drawdowns integrated over time. It shows for how many days a particular drawdown was experienced by the basket of currencies. Notice that the two dynamic hedging methodologies avoid drawdowns that are greater than 6% (these are daily returns) and therefore are more optimal from a drawdown perspective. The 6% bin indicates that the iVaR portfolio was a few days less in stress than the minimum variance portfolio. One might actually argue that this difference is negligible. Table 4.2 confirms that the dynamic hedging methodologies are comparable, especially in terms of their hedging effectiveness and drawdown reduction. Hansen's Model Condifidence set reaffirms that the dynamic hedges are in terms of variance, Sharpe ratio, average drawdowns, Calmar ratio and pain ratio not statistically different from each other.

Let's dig a bit deeper in the analysis of the behaviour of both the minimum variance optimiser and the iVaR optmiser. iVaR's optimal weights on 2006-01-31 are:

`{INR: 0.0102, SEK: 0.0104, USD: 0.0282, RUB: 0.0006, ZAR: 0.0088, KRW: 0.0156, BRL: 0.0101, AUD: 0.0212, GBP: 0.0814, JPY: 0.1162, CAD: 0.0335, HKD: 0.0230,`

TWD: 0.0118, CHF: 0.0277, EUR: 0.600}

while the given bounds on each single currency are

{INR: 0.0102, SEK: 0.0104, USD: 0.4212, BRL: 0.0101, ZAR: 0.0088, AUD: 0.0212, GBP: 0.0814, CHF: 0.0277, HKD: 0.0230, CNY: 0.0, RUB: 0.0006, KRW: 0.0156, SAR: 0.0, JPY: 0.122, CAD: 0.0335, TWD: 0.011}

The currencies INR, SEK, ZAR, AUD, GBP, BRL, KRW, HKD, CHF and almost JPY are set equal to their upper bounds. The above weights are now compared to the ones obtianed from the minimum variance portfolio:

{AUD: 0.0212, GBP: 0.0814, CHF: 0.0277, USD: 0.0811, RUB: 0.0006, INR: 0.0015, BRL: 0.0101, CAD: 0.0335, JPY: 0.1224, HKD: 0.0001, ZAR: 0.0088, SEK: 0.0104, EUR: 0.6}.

Also for this optimisation problem AUD, GBP, CHF, BRL, CAD, SEK, ZAR and almost JPY are set exactly equal to their upperbounds. Both optimisation methodologies are in fact selling USD since it seems to have both a high pair-wise correlation and a high pair-wise coiVaR with other assets. Another interesting relationship can be deduced from the above example. The lower the EUR exposure gets, the more both optimisers have to select less optimal currencies in terms of correlation coefficients and coiVaR and the more alike both porfolios become. The converse is also true, that is the case where the EUR exposure rises due to for instance a higher hedge ratio. Consider for example the bound set on BRL= 0.0101, it is currently only a fraction $\frac{0.0101}{0.4}$ of the total basket of currencies. In case the exposure of Euro rises, the fraction of BRL will become larger. It seems counter-intuitive, but increasing the hedging ratio will benefit the optimisation problem in the methodologies' effectiveness.

The question still remains: Is the optimal iVaR optimiser able to perform the assignment what it is designed for? Appendix B.2 presents the time series plot and the cumulative histogram for the base case when the bounds are relaxed. Notice that this is equivalent to increasing the hedging ratio with respect to the bounds. It is now possible to assign a weight up to tenfold of the current exposure. The time series of the iVaR portfolio is clearly more monotone. The histogram shows that the iVaR portfolio is able to avoid extensive drawdowns and additionally experiences fewer days of medium drawdowns compared to the min variance portfolio. This is also shown in the table (C.2) listed in appendix C.2. Lastly, iVaR portfolio performs better on both the HE and DR index.

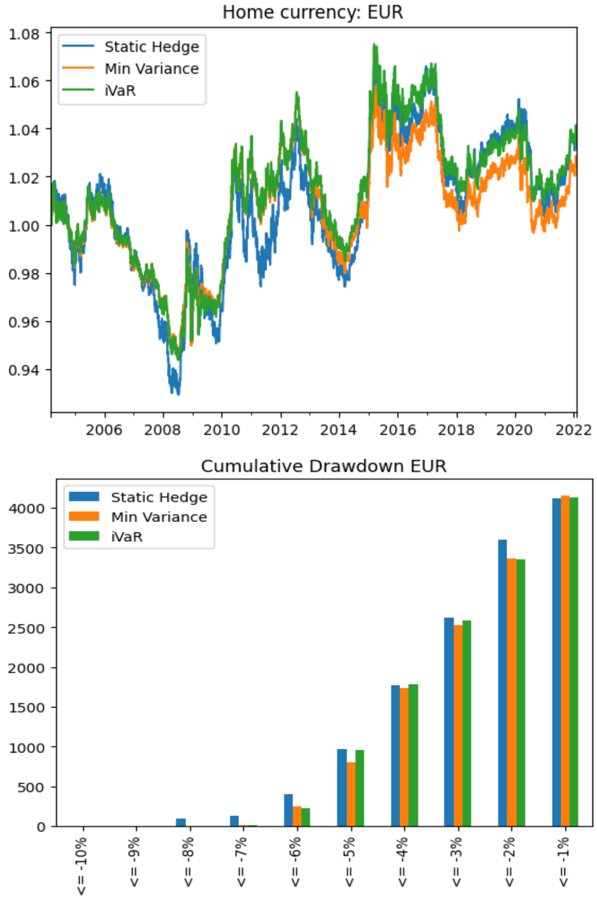


Figure 4.5: Out of sample results for the benchmark, minimum variance and iVaR portfolio.

Annualised: Base case

	Benchmark	**Hist Var**	**iVAR**
Returns	0.0024	0.0016	0.0023
SD	0.0295	0.0249	0.0251
Sharpe Ratio	0.0830	0.0680	0.0939
Avg Drawdown	0.0343	0.0324	0.0328
Max Drawdown	0.09	0.071	0.0724
Calmar ratio	0.0272	0.0237	0.0325
Pain ratio	0.0715	0.0523	0.0717
HE	/	0.1561	0.1511
DR	/	0.055	0.0430

Table 4.2: Evaluation criteria for EUR portfolios.

4.5.2 Changing from home currency

The results from previous section are not case specific for only Euro investors, but generalise more broadly to different home currencies. Figure (4.6) demonstrates the time series for the static benchmark, the iVaR and the historical minimum variance porfolio with USD and GBP as home currencies. The corresponding drawdown histograms are shown in figure (4.7). By visual inspection, one can conclude that the iVaR portfolio minimises the extreme cumulative drawdowns in the case the base currency is GBP, while it is less clear for USD though. The explanation is in fact more subtle and inherent to the optimisation problem. First and foremost, the USD takes already a material fraction of the basket's exposure, i.e. $[40\%, 50]$. By selling 50% of the FX exposure, the basket might already have a position of up to 75% in USD. This implies that there is less freedom for the optimiser to choose among the universe of currencies.

Since it is only possible to hold 100% exposure, diversification possibilities in terms of coiVaR, see appendix (A.2), are limited. For instance, iVaR's optimal weights on 2006-01-31 are:

{TWD: 0.0118, JPY: 0.0810, CAD: 0.0335, KRW: 0.0156, INR: 0.0102, AUD: 0.0091, ZAR: 0.0042, HKD: 0.023, RUB: 0.0006, BRL: 0.0101, GBP: 0.0647, USD: 0.735}

while the imposed bounds are:

{SEK: 0.0104, CNY: 0.0, JPY: 0.122, KRW: 0.0156, ZAR: 0.008, SAR: 0.0, RUB: 0.0006, HKD: 0.023, CHF: 0.027, TWD: 0.0118, 'CAD: 0.0335, INR: 0.0102, EUR: 0.1512, AUD: 0.0212, BRL: 0.0101, GBP: 0.0814}

It is clear that the iVaR optimiser immediately assigns a weight of 100% to the currencies that have a low coiVaR measure, that is CAD, INR, KRW and BRL. Also notice that TWD and HKD are pegged to the USD and hence this bakset has even a larger dollar exposure. This is ideal for an investor whose base currency is USD. In sum, only more or less 25% can be assigned to a set of low coiVaR currencies which makes this optimisation extremely limited.

{RUB: 0.0006, HKD: 0.0230, KRW: 0.0156, JPY: 0.0777, BRL: 0.0101, CAD: 0.0335, TWD: 0.0118, INR: 0.01023, GBP: 0.0814, USD: 0.7357}

Both the iVaR and minimum variance assign maximum weights to the pegged dollar currencies. In addition, they also select the maximum of BRL, CAD, INR and KRW while the difference in JPY is small. This once again explains why both dynamic portfolio's are so alike.

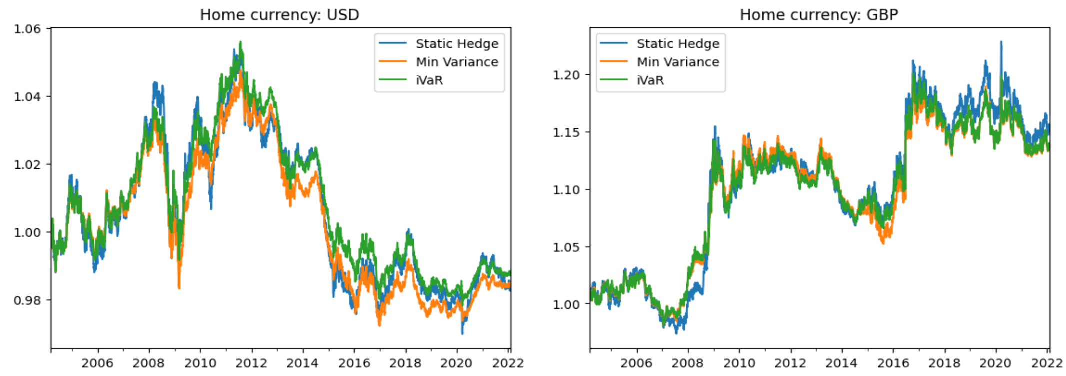


Figure 4.6: Time series for the benchmark, minimum variance and iVaR portfolio denominated in USD and GBP.

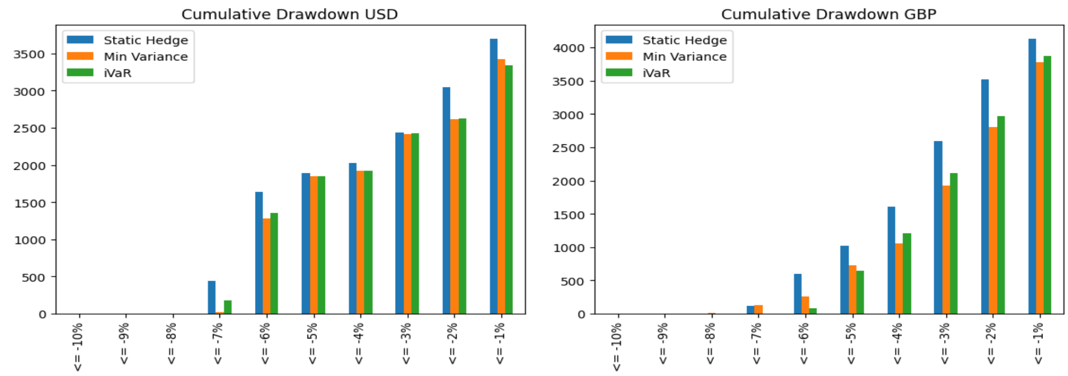


Figure 4.7: Out of sample realised drawdowns for the benchmark, minimum variance and iVaR portfolio denominated in USD and GBP.

Now lets take a closer look at the GBP optimisation problem which has only an a varying exposure of $[5\%, 10\%]$ in the MSCI ACWI. This problem has to be free in nature compared to the USD case, which should therefore favor diversification opportunities. This phenomenon is indeed observed in the drawdown plot (figure 4.7).

Appendix (B.3) and **??** list also the plots and tables for the unconstrained optimisation problem. The USD denominated iVaR portfolio now has fewer extreme drawdowns and scores well on the HE and DR indices. This section concludes that the portfolio's home currency does not influence the previous findings. The iVaR is in general only marginally better in terms of thee time length that extreme drawdowns occur with strict bounds. If these bounds are to be relaxed, iVaR portfolios clearly score better on the cumulative drawdowns plots. It is also evidenced by the drawdown reduction index.

Annualised: USD			
	Benchmark	**Hist Var**	**iVaR**
Returns	-0.0008	-0.0008	-0.0006
SD	0.0154	0.01159	0.0117
Sharpe Ratio	-0.0508	-0.0691	-0.0520
Avg DD	0.0370	0.0337	0.03411
Max DD	0.0806	0.0721	0.0738
Calmar ratio	-0.0097	-0.0111	-0.0082
Pain ratio	-0.0211	-0.0237	-0.0179
HE	/	0.2494	0.2386
DR	/	0.0891	0.0789

Table 4.3: Evaluation criteria for USD portfolios.

Annualised: GBP			
	Benchmark	**Hist Var**	**iVaR**
Returns	0.0083	0.0074	0.0075
SD	0.03493341	0.0317	0.03193651
Sharpe Ratio	0.23976125	0.2353	0.235698803
Avg DD	0.03393752	0.0276	0.028009615
Max DD	0.07604856	0.082732436	0.067112372
Calmar ratio	0.11013592	0.090297911	0.112161096
Pain ratio	0.24679699	0.269988235	0.268743331
HE	/	0.091518862	0.085788985
DR	/	0.184679183	0.174671216

Table 4.4: Evaluation criteria for GBP portfolios.

4.5.3 The devil is in transaction costs, or not?

This research considers transaction costs in order to put the obtained results in a more realistic setting. The time series and drawdown histogram are displayed in figure (4.8) and the statistics are summarised in table (4.5). We notice that the performance of the dynamic hedges deteriorates from the moment transaction costs are taken into account. Nevertheless, the iVaR portfolio is still able to reduce the average drawdown compared to the static hedge, which isn't the case for the minimum variance portfolio. The latter has even a negative value for the DR index while the HE index is similar to iVaR's score. In terms of drawdowns and the realised volatility, the iVaR portfolio is more robust while the minimum variance portfolio is only stable for the realised volatility.

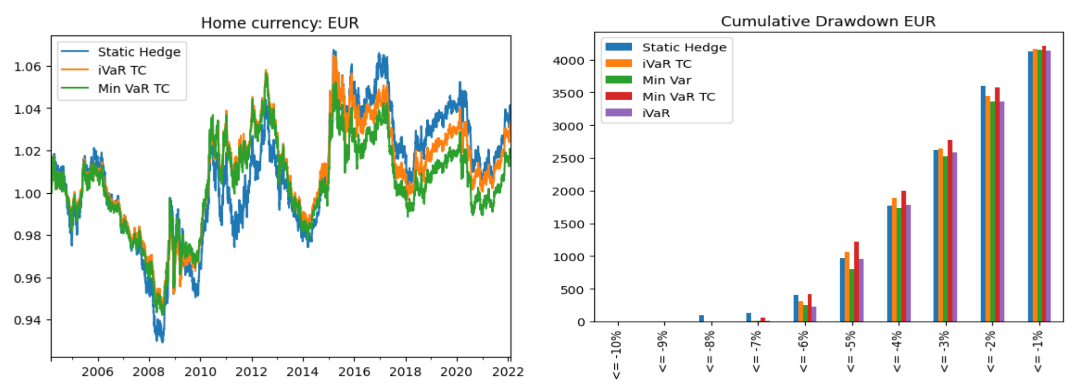


Figure 4.8: test

Annualised: tc					
	Benchmark	**Hist Var**	**iVaR**	**Hist Var TC**	**iVaR TC**
Returns	0.002457	0.00169762	0.00235764	0.00128829	0.00190195
SD	0.02957021	0.02495331	0.02510144	0.0251839	0.02525435
Sharpe Ratio	0.08309039	0.06803179	0.09392465	0.05115537	0.07531162
Avg DD	0.03432238	0.0324086	0.03284459	0.0354002	0.03389938
Max DD	0.09006135	0.07146387	0.07240098	0.07442263	0.07242134
Calmar ratio	0.02728141	0.02375492	0.03256371	0.01731049	0.02626223
Pain ratio	0.07158595	0.05238171	0.0717818	0.03639222	0.05610562
HE	/	0.15613361	0.15112412	**0.14833543**	0.14595314
DR	/	0.05575892	0.04305616	-0.0314027	**0.01232444**

Table 4.5: Evaluation criteria for EUR portfolios where transaction costs are considered.

4.5.4 What about different covariance estimation methods?

Using the acquired knowledge of previous results, we would infer that the effect of using different covariance estimators will not be material. Table (C.1) endorses our intuition although one estimator could be favored, i.e. the L1-shrinkage method. It is able to reduce 15% of the incurred volatility and 6% of the average drawdowns. The DCC and CCC estimators seem to inhibit a similar behaviour since they obtain the highest return and Sharpe ratio, however DCC scores better on the DR index. A visual inspection of the cumulative drawdown plot leads to the conclusion that iVaR, L1-shrinkage and the traditional unbiased estimator of the covariance matrix ought to be preferred when the goal is to minimise the cumulative drawdowns, especially for large losses.

In the event the constraint are abolished, iVaR is among the better performers based on the combination HE and DR, but achieves the best scores on the Calmar and Pain ratio. The EWMA covariance estimator performs in fact better than expected and is able score 32% on the HE index. One can clearly see that its time series is the smoothest among its candidates. The L-1 estimator achieves the best score on the hedging effective index, which is a clear pattern.

Annualised								
	Benchmark	**Hist Var**	**iVaR**	**DCC**	**CCC**	**L1**	**L2**	**EWMA**
Returns	0.002457	0.00169762	0.00235764	0.0025977	**0.0026457**	0.00167011	0.00052707	0.00089474
SD	0.02957021	0.02495331	0.02510144	0.02604051	0.02619257	**0.02493744**	0.02517965	0.02571186
Sharpe Ratio	0.08309039	0.06803179	0.09392465	0.09975609	**0.1010095**	0.06697194	0.02093238	0.03479886
Avg DD	0.03432238	0.0324086	0.03284459	0.03422567	0.03496565	**0.0323849**	0.03370924	0.03481611
Max DD	0.09006135	0.07146387	0.07240098	0.07664934	0.07661399	**0.07132955**	0.07815689	0.07292899
Calmar ratio	0.02728141	0.02375492	0.03256371	0.03389069	**0.03453284**	0.02341398	0.00674374	0.01226869
Pain ratio	0.07158595	0.05238171	0.0717818	**0.07589914**	0.07566566	0.05157058	0.01563577	0.02569912
HE	/	0.15613361	0.15112412	0.11936693	0.11422432	**0.15667037**	0.14847905	0.13048098
DR	/	0.05575892	0.04305616	0.00281771	-0.0187419	**0.05644941**	0.01786431	-0.0143849

Table 4.6:

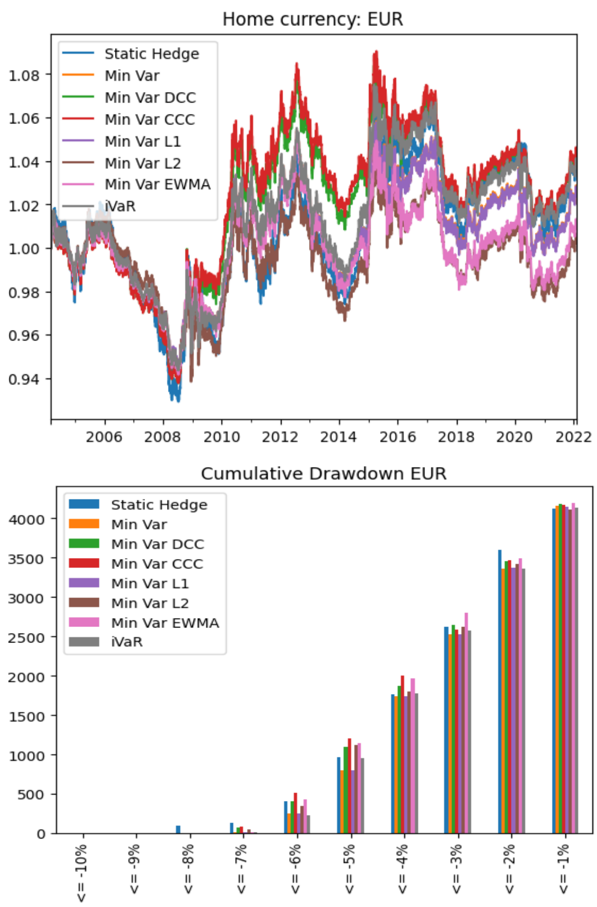


Figure 4.9: test

Chapter 5

Conclusion

This paper examined different forex hedging case studies in order to determine the effectiveness of an iVaR optimiser. First of all, the results in this paper suggest that hedging FX risk under different methodologies in an restricted setting give rise to similar results in terms of hedging effectiveness and drawdown reduction. Nevertheless, the dynamic hedges are effective in beating a static hedge based on these performance criteria. From the moment one relaxes some strict bounds, the different methodologies increase in their effectiveness which is a clear sign of the complexity of the problem one encounters in determining the optimal basket of currencies. This relaxation of the problem can be suble as well. For instance, by increasing the hedge ratio we will naturally converge to an unconstrained optimisation problem. In the limit, however, our portfolio will not have any material exposure anymore. Another empirical observation made in this paper, is that home currencies which start with a low exposure are more free in nature with respect to the optmisation problem and therefore facilitating hedging effectiveness. These results are obtained in multiple backtests: using different home currencies, incorporating transaction costs and using state of the art covariance estimators.

Further research on this topic is however advised. First and foremost, one could conduct an experiment in considering this framework for the universe of stocks of MSCI ACWI together with the universe of foreign currencies. Next, iVaR is only using historical paths which is as such only one realisation of a stochastic process. It is therefore advised to better learn the inherent stochastic process such that a simulation approaches could be used.

Bibliography

Baillie, Richard T, Robert J Myers. 1991. Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics* **6**(2) 109–124. Publisher: Wiley Online Library.

Beine, Michel, Charles S. Bos, Sébastien Laurent. 2006. The Impact of Central Bank FX Interventions on Currency Components. *Journal of Financial Econometrics* **5**(1) 154–183.

Black, Fischer, Myron Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 637–659.

Bollerslev, Tim. 1990. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *The review of economics and statistics* **72**(3) 498–505. Publisher: JSTOR.

C Fishburn, Peter. 1977. Mean-risk analysis with risk associated with below- target returns. *The American Economic Review* **67**(2) 116–126.

Caccioli, Fabio, Imre Kondor, Matteo Marsili, Susanne Still. 2016. Liquidity risk and instabilities in portfolio optimization. *International Journal of Theoretical and Applied Finance* **19**(05). Publisher: World Scientific.

Cecchetti, Stephen G, Robert E Cumby, Stephen Figlewski. 1988. Estimation of the optimal futures hedge. *The Review of Economics and Statistics* 623–630Publisher: JSTOR.

Chekhlov, Alexei, Stanislav Uryasev, Michael Zabarankin. 2005. Drawdown measure in portfolio optimization. *International Journal of Theoretical and Applied Finance* **8**(01) 13–58.

Choudhry, Taufiq. 2003. Short-run deviations and optimal hedge ratio: evidence from stock futures. *Journal of Multinational Financial Management* **13**(2) 171–192. Publisher: Elsevier.

Chowdhury, Abdur R. 1991. A causal analysis of defense spending and economic growth. *Journal of Conflict Resolution* **35**(1) 80–97. Publisher: Sage Publications 2455 Teller Road, Newbury Park, CA 91320.

Crowder, William J, Anas Hamed. 1993. A cointegration test for oil futures market efficiency. *Journal of Futures Markets* **13** 933–933. Publisher: JOHN WILEY & SONS LTD.

Dwyer Jr, Gerald P, Myles S Wallace. 1992. Cointegration and market efficiency. *Journal of International Money and Finance* **11**(4) 318–327. Publisher: Elsevier.

Engle, Robert. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* **20**(3) 339–350.

Fan, Jianqing, Runze Li. 2001. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association* **96**(456) 1348–1360. Publisher: Taylor & Francis.

Gagnon, Louis, Gregory J Lypny, Thomas H McCurdy. 1998. Hedging foreign currency portfolios. *Journal of Empirical Finance* **5**(3) 197–220. Publisher: Elsevier.

Hansen, Peter R, Asger Lunde, James M Nason. 2011. The model confidence set. *Econometrica* **79**(2) 453–497. Publisher: Wiley Online Library.

Johnson, Leland L. 1976. The theory of hedging and speculation in commodity futures. *The Economics of Futures Trading*. Springer, 83–99.

Jondeau, Eric, Michael Rockinger. 2005. Conditional asset allocation under non-normality: how costly is the mean-variance criterion? *European Financial Management* .

Kroner, Kenneth F, Jahangir Sultan. 1993. Time-varying distributions and dynamic hedging with foreign currency futures. *Journal of financial and quantitative analysis* **28**(4) 535–551. Publisher: Cambridge University Press.

Ledoit, Olivier, Michael Wolf. 2003. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance* **10**(5) 603–621. Publisher: Elsevier.

Lee, Hsiang-Tai, Jonathan Yoder. 2007. Optimal hedging with a regime-switching time-varying correlation GARCH model. *Journal of Futures Markets* **27**(5) 495–516. Publisher: Wiley Online Library.

Markowitz, Harry. 1952. The utility of wealth. *Journal of political Economy* **60**(2) 151–158.

Markowitz, Harry. 1959. *Portfolio Selection: Efficient Diversification of Investments.*. Yale University Press.

Myers, Robert J. 1991. Estimating time-varying optimal hedge ratios on futures markets. *The Journal of Futures Markets* **11**(1) 39. Publisher: Wiley Periodicals Inc.

Nawrocki, David N. 1999. A brief history of downside risk measures. *The Journal of Investing* **8**(3) 9–25.

Ozturk, Murat, Ismail Hakki Toroslu, Guven Fidan. 2016. Heuristic based trading system on Forex data using technical indicator rules. *Applied Soft Computing* **43** 170–186.

Park, Tae H, Lorne N Switzer. 1995. Time-varying distributions and the optimal hedge ratios for stock index futures. *Applied Financial Economics* **5**(3) 131–137. Publisher: Taylor & Francis.

Petropoulos, Anastasios, Sotirios P. Chatzis, Vasilis Siakoulis, Nikos Vlachogiannakis. 2017. A stacked generalization system for automated FOREX portfolio trading. *Expert Systems with Applications* **90** 290–302.

Roy, Andrew Donald. 1952. Safety first and the holding of assets. *Econometrica: Journal of the econometric society* 431–449.

Tong, Wilson HS. 1996. An examination of dynamic hedging. *Journal of International Money and Finance* **15**(1) 19–35. Publisher: Elsevier.

Égert, Balázs, Evžen Kočenda. 2014. The impact of macro news and central bank communication on emerging European forex markets. *Economic Systems* **38**(1) 73–88.

Appendix A

Mathematical tools

A.1 Maximum likelihood estimation

The parameters of the univariate GARCH can be estimated by the maximum likelihood estimation. In particular, when considering a standard normal random variable at time t $z_t \sim \mathcal{N}(0,1)$, which is discrete in time. Then by the linear transformation of a standard normal random variables we obtain $a_t = \sigma_t z_t$ which is also normally distributed. The likelihood function for an univariate GARCH(p,q) model with parameters c, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is then

$$\begin{aligned} \mathrm{L}(c, \boldsymbol{\alpha}, \boldsymbol{\beta} | a_{m+1}, ..., a_T) &= f(a_{m+1}, ..., a_T | c, \boldsymbol{\alpha}, \boldsymbol{\beta}, a_1, ..., a_m) \\ &= \prod_{t=(m+1)}^{T} \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left\{ \frac{-a_t^2}{2\sigma_t^2} \right\}. \end{aligned} \tag{A.1}$$

First notice that not the entire sample size is used but only observations $m+1, ..., T$, where $m = \max(p,q)$. That is because the first m observations are necessary to produce an estimate for a_{m+1}. The likelihood function in equation (A.1) is in fact a conditional likelihood function. Also, the variance estimates σ_t^2 have to be calculated recursively

$$\sigma_t = c + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \quad t = m+1, ..., T. \tag{A.2}$$

Since $\ln(x)$ is a strictly monotone increasing function, maximising the log-likelihood instead of the likelihood, is equivalent.

$$\begin{aligned}\ln(\mathrm{L}) &= \ln(\mathrm{L}(c, \boldsymbol{\alpha}, \boldsymbol{\beta}|a_{m+1}, ..., a_T) \\ &= l(a_{m+1}, ..., a_T|c, \boldsymbol{\alpha}, \boldsymbol{\beta}, a_1, ..., a_m) \\ &= -\frac{1}{2}\sum_{t=m+1}^{T}\left(\ln(2\pi) + \ln(\sigma_t^2) + \frac{a_t^2}{\sigma_t^2}\right) \\ &= \text{constant} - \frac{1}{2}\sum_{t=m+1}^{T}\left(\ln(\sigma_t^2) + \frac{a_t^2}{\sigma_t^2}\right).\end{aligned} \tag{A.3}$$

A.2 Co - iVaR matrix

The correlation matrix is an useful tool to summarise the degree of linear dependency among random variables. It is paramount for the minimum variance optimisation, however, not quite useful for data that do not inhibit a Gaussian dependency structure. If we for example want to minimise the cumulative drawdown, which implicitly also concerns tail behaviour, co-iVaR of two random variables needs to be measured instead. Co-iVaR is a pairwise association measure between assets which compares the risk of an equally weighted portfolio of asset X and Y with the risk of holding X and Y individually:

$$\mathrm{coiVaR}(X, Y) = \frac{\mathrm{iVaR}(X + Y)}{\mathrm{iVaR}(X) + \mathrm{iVaR}(Y)}. \tag{A.4}$$

The above formula is an elegant way of measuring the potential pairwise diversification benefits of using two assets in a portfolio. Drawdown risk measures are sub-additive and non-negative, thus

$$0 \leq \mathrm{coiVaR}(X, Y) \leq 1. \tag{A.5}$$

In other words, adding a financial instrument to a portfolio where all the pairwise coiVaR's are low, will not increase the portfolio's iVaR. In the extreme case where $\mathrm{coiVaR}(X, Y) \sim 0$, the combination of the time series of asset X and Y will be close to monotonic growth while the separate time series are not necessarily close to monotonic growth. The equally weighted drawdowns for the portfolio of these two assets in severity, frequency and integrated over time (time to earn back), is low.

Finally, it is possible to construct a co-iVaR matrix that is comparable to the correlation matrix. It is an elegant way of analysing the pairwise correlation in drawdowns

$$
\boldsymbol{C} = \begin{pmatrix} 1 & \mathrm{coiVaR}(X_1, X_2) & \dots & \mathrm{coiVaR}(X_1, X_N) \\ \mathrm{coiVaR}(X_2, X_1) & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & \vdots \\ \mathrm{coiVaR}(X_N, X_1) & \dots & \mathrm{coiVaR}(X_N, X_{N-1}) & 1 \end{pmatrix}. \tag{A.6}
$$

Appendix B

Figures

B.1 Supporting figures for descriptive statistics

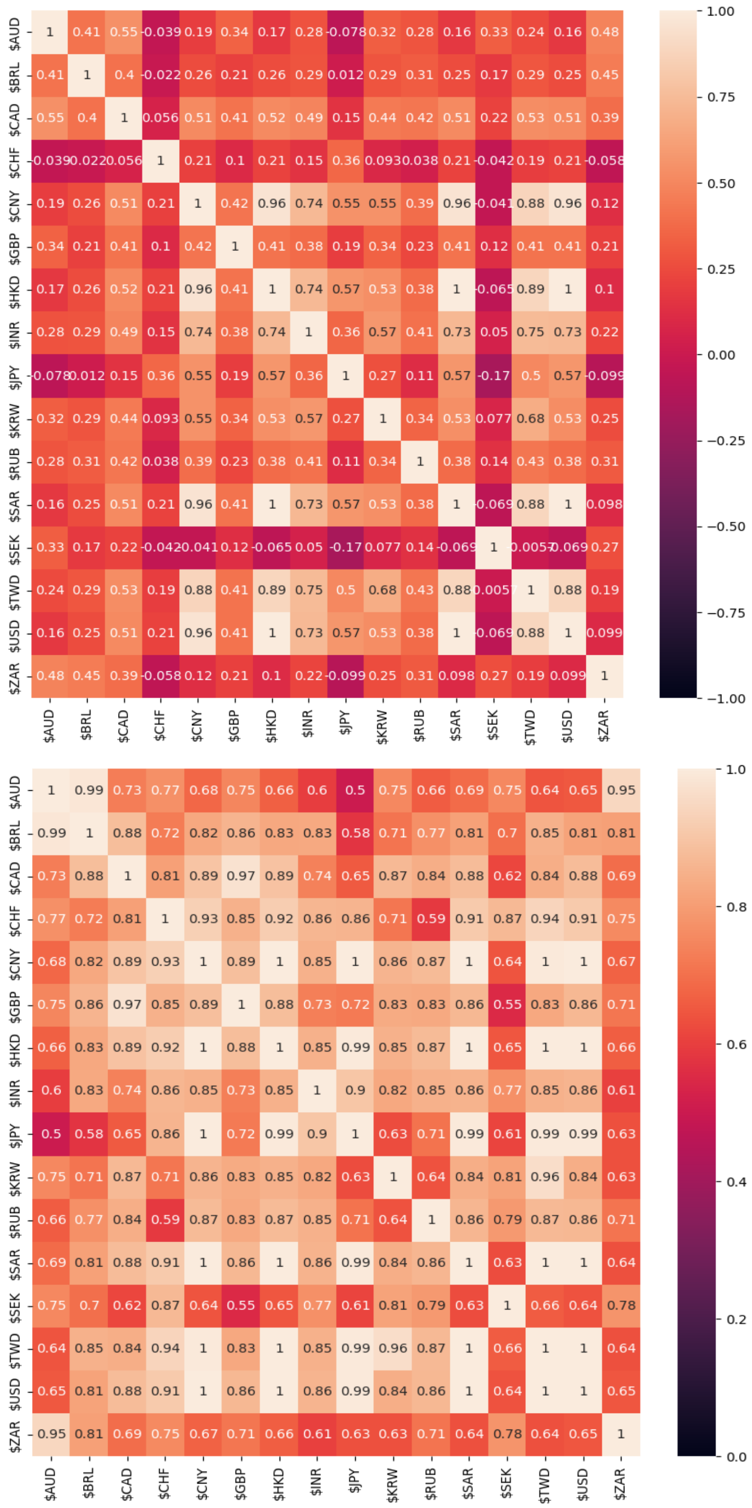


Figure B.1: The correlation matrix and co-iVaR matrix for the full backtest period: January 2004 - January 2022.

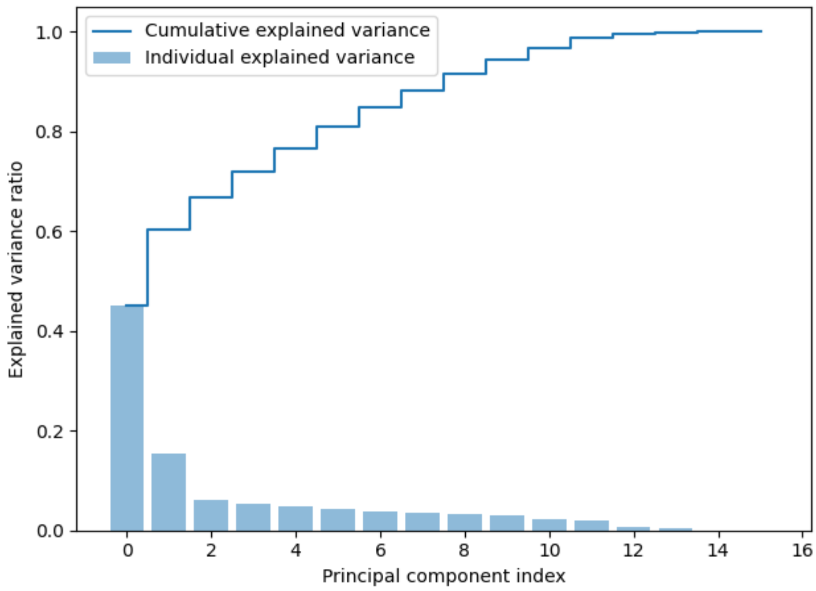


Figure B.2: The explained variance plot for the correlation matrix of January 2004 - January 2022.

B.2 Supporting figures for results: section (4.5.1)

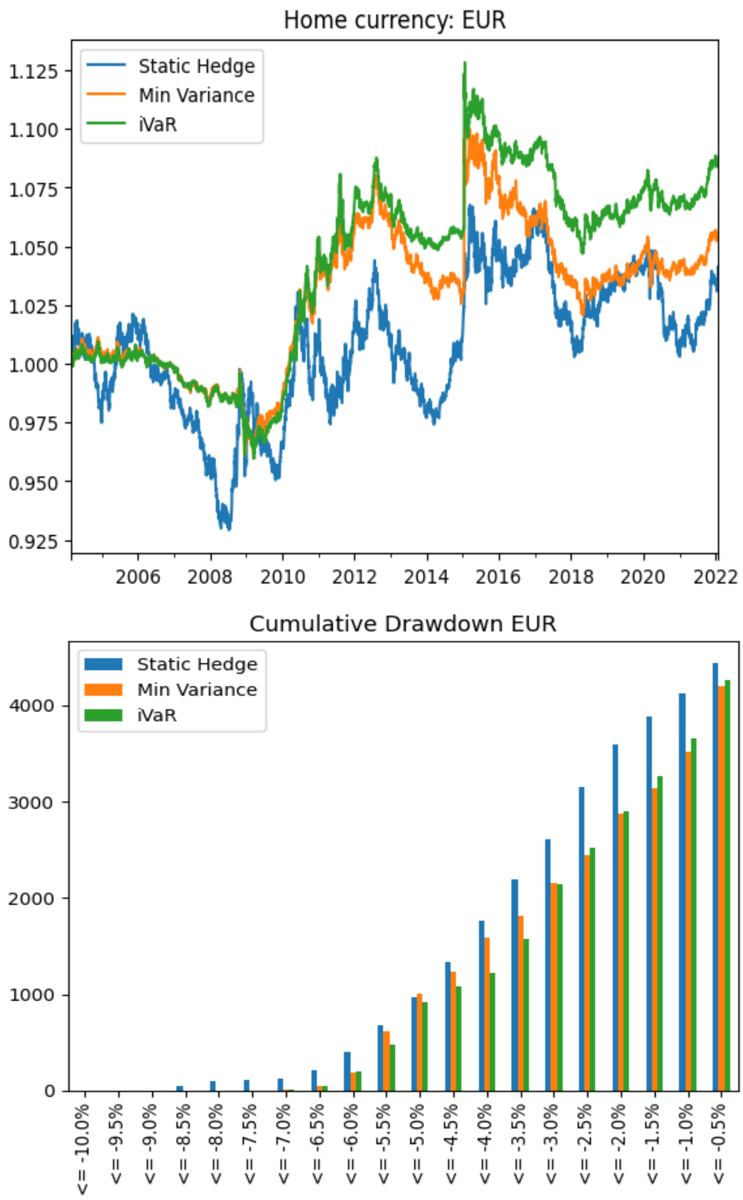


Figure B.3: Unconstrained optimistation problem for base case.

B.3 Supporting figures for results: section (4.5.2)

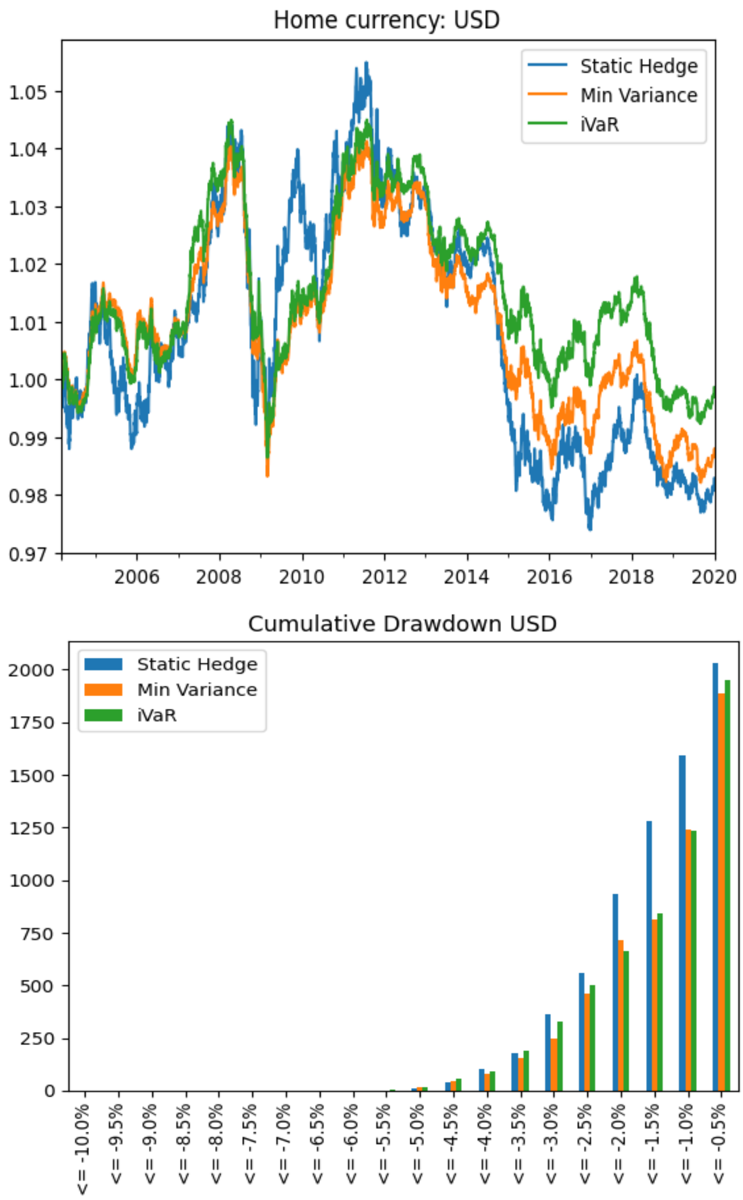
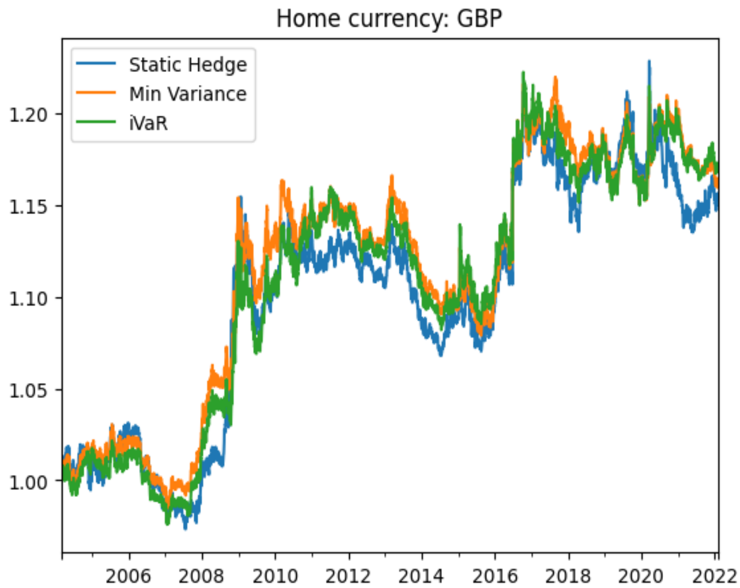


Figure B.4: Unconstrained optimistation problem for USD denominated portfolio.



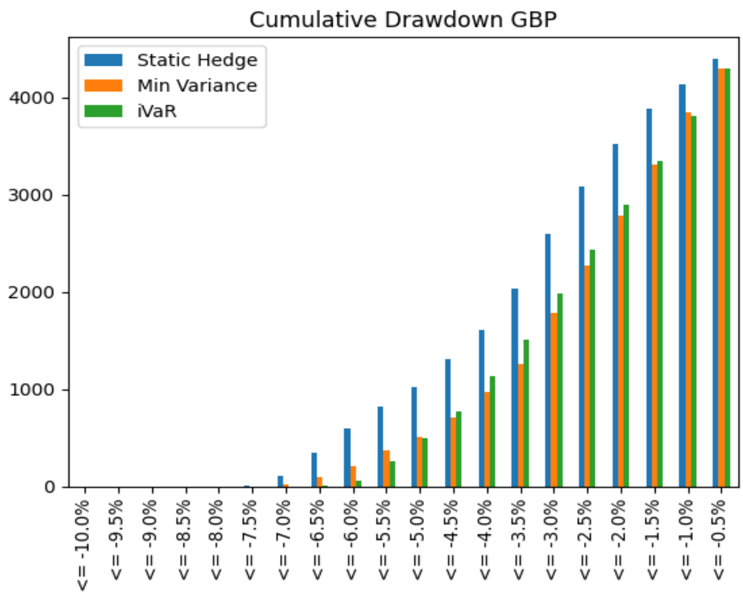


Figure B.5: Unconstrained optimistation problem for USD denominated portfolio.

B.4 Supporting figures for results: section (4.5.4)

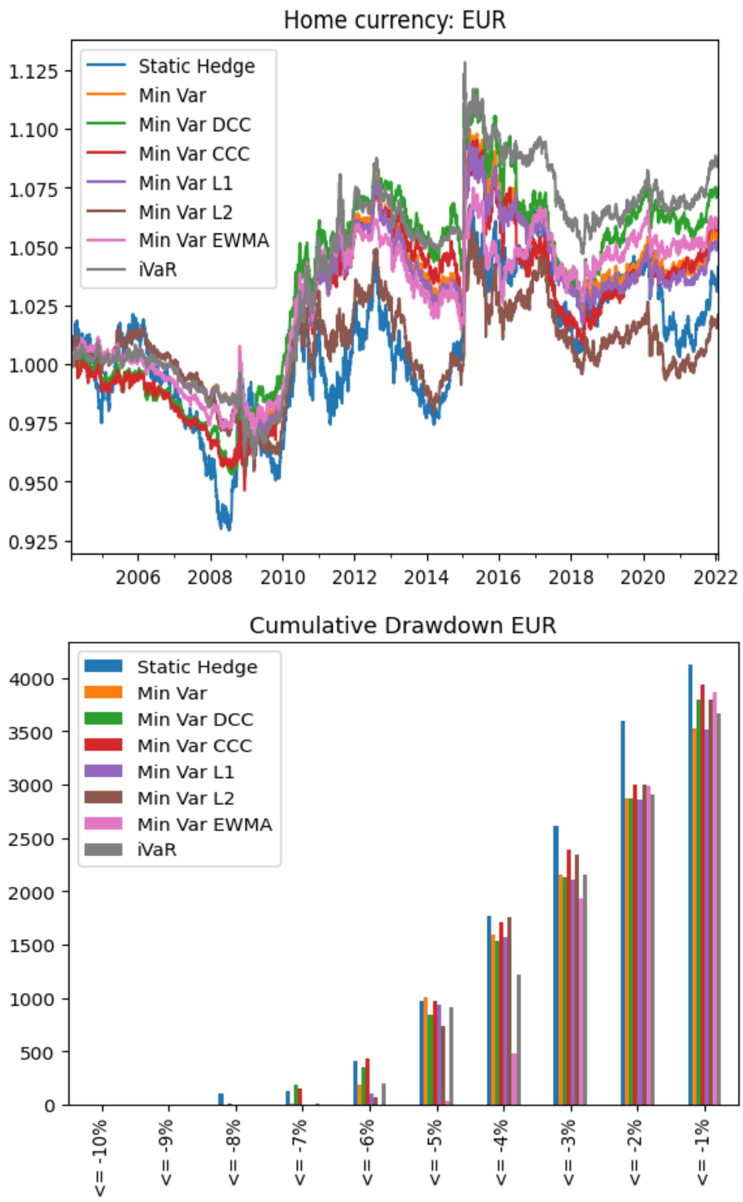


Figure B.6: Unconstrained optimistation problem for different covariance estimators.

Appendix C

Tables

C.1 Table supporting data description

Symbol	**Currency name**
AED	Emirati Dirham
ARS	Argentine Peso
AUD	Australian Dollar
BRL	Brazilian Real
CAD	Canadian Dollar
CHF	Swiss Franc
CLP	Chilean Peso
CNY	Chinese Yuan
COP	Colombian Peso
CZK	Czech Krone
DKK	Danish Krone
EGP	Egyptian Pound
EUR	EURO
GBP	British Pound Sterling
HKD	Hong Kong Dollar
HUF	Hungarian Forint
IDR	Indonesian Rupiah
ILS	Israeli New Shekel
INR	Indian Rupee
JOD	Jordanian Dinar
JPY	Japanese Yen
KRW	Korean Won
KWD	Kuwaiti Dinar
MAD	Moroccan Dirham
MXN	Mexican Peso
MYR	Malaysian Ringgit
NOK	Norwegian Krone
NZD	New Zealand Dollar
PEN	Peruvian Sol
PHP	Philippine Peso
PKR	Pakistani Rupee
PLN	Polish Zloty
QAR	Qatari Riyal
RUB	Russian Ruble
SAR	Saudi Riyal
SEK	Swedish Krona
SGD	Singapore Dollar
THB	Thai Baht
TRY	Turkish Lira
TWD	New Taiwan Dollar
USD	United States Dollar
VES	Venezuelan Bolívar
ZAR	South African Rand

Table C.1: The entire list of currencies where at some point during the backtest period stocks where denominated in for MSCI ACWI.

C.2 Tables supporting results

Annualised: EUR			
	Benchmark	**Hist Var**	**iVAR**
Returns	0.002457	0.00304821	0.00462406
SD	0.02957021	0.02235569	0.02195305
Sharpe Ratio	0.08309039	0.13635046	0.21063422
Avg Drawdown	0.03432238	0.028921	0.02847178
Max Drawdown	0.09006135	0.07224495	0.07196252
Calmar ratio	0.0272814	0.04219269	0.06425656
Pain ratio	0.07158595	0.10539778	0.16240868
HE	/	0.24397924	0.25759568
DR	/	0.15737204	0.17046027

Table C.2:

Annualised: USD			
	Benchmark	**Hist Var**	**iVaR**
Returns	-0.0007849	-0.0006396	-0.0001107
SD	0.01544462	0.00885366	0.00896374
Sharpe Ratio	-0.0508215	-0.0722451	-0.012354
Avg DD	0.03703833	0.02744747	0.02488278
Max DD	0.0806016	0.05660403	0.05595632
Calmar ratio	-0.0097383	-0.0113001	-0.001979
Pain ratio	-0.0211921	-0.0233039	-0.0044504
HE	\	0.42674774	0.41962052
DR	\	0.25894413	0.32818819

Table C.3:

Annualised: GBP			
	Benchmark	**Hist Var**	**iVAR**
Returns	0.00837568	0.00858326	0.00901704
SD	0.03493341	0.03176003	0.03288017
Sharpe Ratio	0.23976125	0.27025346	0.27423959
Avg DD	0.03393752	0.0263928	0.02699122
Max DD	0.07604856	0.07490787	0.06721159
Calmar ratio	0.11013592	0.11458422	0.13415906
Pain ratio	0.24679699	0.32521211	0.33407327
HE	/	0.09084076	0.05877587
DR	/	0.22231206	0.20467921

Table C.4:

Annualised: EUR								
	Benchmark	**Hist Var**	**iVaR**	**DCC**	**CCC**	**L1**	**L2**	**EWMA**
Returns	0.002457	0.00304821	**0.00462406**	0.00402016	0.00328554	0.00282008	0.00125537	0.0033017
SD	0.02957021	0.02235569	0.02195305	0.02455679	0.0240269	**0.0217366**	0.02266926	0.02258858
Sharpe Ratio	0.08309039	0.13635046	**0.21063422**	0.16370889	0.13674412	0.12973859	0.05537754	0.14616691
Avg DD	0.03432238	0.028921	0.02847178	0.02999093	0.03158229	0.02846646	0.02983313	**0.02478582**
Max DD	0.09006135	0.07224495	0.07196252	0.08171103	0.07955409	0.0703685	0.06644364	**0.05333731**
Calmar ratio	0.02728141	0.04219269	**0.06425656**	0.04919977	0.04129942	0.04007582	0.01889373	0.06190233
Pain ratio	0.07158595	0.10539778	**0.16240868**	0.13404599	0.10403099	0.09906659	0.04207966	0.1332094
HE	/	0.24397924	0.25759568	0.16954313	0.18746267	**0.26491573**	0.23337505	0.2361034
DR	/	0.15737204	0.202296	0.15213142	0.0913639	0.18541778	0.15770333	**0.31966368**

Table C.5: