BI

WISEflow
Europe/Oslo(CEST)
30 Jun 2022

Handelshøyskolen BI

GRA 19703 Master Thesis

Thesis Master of Science 100% - W

Predefinert informasjon

Startdato:	16-01-2022 09:00	**Termin:**	202210
Sluttdato:	01-07-2022 12:00	**Vurderingsform:**	Norsk 6-trinns skala (A-F)
Eksamensform:	T		
Flowkode:	202210||10936||IN00||W||T		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:	Daria Sokolova og Elitsa Todorova Boyadzhieva

Informasjon fra deltaker

Tittel *:	Portfolio risk hedging: Currency exposure and Shrinkage techniques
Navn på veileder *:	Bruno Gerard

Inneholder besvarelsen konfidensielt materiale?:	Nei	**Kan besvarelsen offentliggjøres?:**	Ja

Gruppe

Gruppenavn:	(Anonymisert)
Gruppenummer:	72
Andre medlemmer i gruppen:	

MSc Thesis

Submitted

in Partial Fulfilment of the Requirements

for the Degree of

MASTERS OF SCIENCE

in

Quantitative Finance

Portfolio risk hedging:
Currency exposure and Shrinkage techniques

Supervisor: Bruno Gerard

Table of Contents

Acknowledgements

This research paper is marking the end of our Master of Science in Quantitative Finance program at BI Norwegian Business School. We are extremely proud to have taken the journey of achieving MSc degree here and we are feeling excited to be so close to calling ourselves "BI alumni".

We would like to give a special thanks to our MSc supervisor Bruno Gerard, for the time, the patience, and his valuable insights! The completion of this paper would have not been possible without his expertise, support, and guidance.

Abstract

We apply different shrinkage techniques to the covariance and concentration matrices used for the minimum variance currency risk hedging of a globally diversified portfolio. The techniques are applied with the aim to induce sparsity in the estimator and by that to reduce the multicollinearity and the noise, resulting in a more robust estimator and decreased out-of-sample portfolio risk. We show that the application of such techniques leads to a worsening of the risk characteristics of the portfolio. We argue that this is likely due to the structure of the minimum variance hedge as well as to the shrinkage methods which seem to disturb the balance and optimality of the minimum variance hedge.

1. Introduction

In 1952 in his paper “Portfolio Selection”, the economist Harry Markowitz presented a rigorous mathematical framework that laid the foundations for modern portfolio theory and portfolio optimization. His work on portfolio selection earned him the Nobel Prize in Economic Sciences in 1990. The paper presented the theory referred to as mean-variance portfolio analysis (MVA) and proposed optimizing portfolios based on the risk and the return of their individual assets holdings while balancing the expected return and the variance of the portfolio. Furthermore, it was shown that for a well-diversified portfolio (when N is large enough), the risk will depend mainly on the pairwise covariances of the assets included in the portfolio rather than their individual variances. In other words, the portfolio risk will be based on the off-diagonal elements of the portfolio´s covariance matrix.

Furthermore, diversification is an essential element in the framework as well as thought to be the only free lunch in finance. There is a large idiosyncratic risk characterizing investments in individual asset classes, industries, as well as countries. The mean-variance framework shows that unsystematic (individual assets) risk can be reduced significantly by simply increasing the variety and number of assets in the portfolio. There are several ways that investors can diversify their portfolios, but one of the most effective is to simply increase the geographical spread of the investments and to diversify globally. Although international markets are connected to some degree and the most significant financial shocks are usually resonating worldwide, correlations between countries are on average lower than correlations within countries (Roll, 1992). For instance, Solnik (1974) shows that the variability of return of an internationally well-diversified portfolio would be one-tenth as risky as typical security and half as risky as a well-diversified portfolio of U.S. stocks (with the same number of holdings). More recent studies, however, such as Eiling & Gerard (2015) find that correlations between emerging and developed equity markets, as well as between emerging market regions, are exhibiting significant upward time trends. Increasing cross country correlations suggest that the benefits of international diversification may be decreasing over time.

Moreover, the reduction of the risk achieved through diversification, does not come at the cost of a reduction of the return as it would if an investor chose to reduce her risk by other means such as investing in short-term government debt (Khoury et al., 1996). However, as De Santis & Gérard (1998) point out, an investment in a foreign asset is a combination of an investment in the performance of the foreign asset and an investment in the performance of the domestic currency relative to the foreign currency. While the reduction of the idiosyncratic risk by international diversification is a well-established fact, the international diversification benefit comes at the cost of introducing additional risk in the portfolio in the form of currency risk. This additional risk is an important element to consider since variations in exchange rates can be large and may have unpredictable effects on the returns of international equities, particularly for countries with unstable monetary policies. Consequently, diversifying internationally may complicate decision-making in portfolio management since it requires making choices on more variables than just the underlying assets. Particularly, investors should make allocation choices based on two factors – the underlying stocks (indexes) and the hedging portfolio of the corresponding currencies.

In theory, the choice of equity selection cannot be done independently of the currency decision (Khoury & Jorion, 1996). However, since many asset managers today lack expertise in currencies, there is a trend in the industry for delegating the currency part of the portfolio over managers specialized in currencies or the so-called "overlay" managers. This allows the "core" asset managers to focus on the equity portfolio where they regularly communicate the core positions to the currency manager. (Khoury & Jorion, 1996)

In this paper we take the position of a currency manager who is handling the currency risk of an internationally diversified portfolio. For simplicity, we use the equally weighted portfolio, but we also test the same models on the minimum variance portfolio, and we achieve similar results. Our aim is to reduce the total risk of the portfolio by optimizing the currency exposure: which currencies and how much of those currencies our internationally diversified investor should hedge in order to optimally minimize the risk of her portfolio.

2. Literature Review

2.1. The mathematical background of the minimum variance hedge

When considering a multivariate portfolio consisting of both foreign and domestic indexes and currencies, the optimal MV portfolio is:

$$\begin{bmatrix} w_N \\ w_{N+1} \end{bmatrix} = \lambda \begin{bmatrix} \Sigma^{-1}\mu \\ 1 - 1'\Sigma^{-1}\mu \end{bmatrix} + (1-\lambda) \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Where:

w_N - positions in N risky assets, both equity indexes and foreign currency bills,

w_{N+1} - position in the riskless asset (the domestic currency bill),

λ - investor's risk tolerance,

Σ^{-1} - the covariance matrix between the indexes and currencies,

μ - the vector of returns of the risky assets.

The equation is standard in the international finance literature and currency positions are contained in both the foreign equity indexes and short-term bills. The domestic bill is the only risk-free asset since it does not contain currency risk as with the foreign bills (Khoury & Jorion, 1996).

However, if we want to evaluate the currency positions individually, we can replace the foreign bills with forward contracts since those can be replicated using foreign and domestic bills (Khoury & Jorion, 1996). Forwards are customizable derivative contracts which makes them particularly useful for hedging. Moreover, they are costless at t = 0 as entering a forward contract involves zero investment and the settlement is done first at the maturity of the contract. Further in the discussion, hedging is referring to entering in a forward contract, which can be done by both short and long positions.

Furthermore, we partition the μ and Σ^{-1} in such a way that they are clearly separated and defined individually:

$$\mu = \begin{bmatrix} \mu_E \\ \mu_C \end{bmatrix} \quad \Sigma = \begin{bmatrix} \Sigma_{EE} & \Sigma_{EC} \\ \Sigma_{CE} & \Sigma_{CC} \end{bmatrix}$$

Where:

Σ – the covariance matrix of the entire portfolio consisting of equities and currencies,

Σ_{EE} - covariance matrix of equity indexes only,

Σ_{EC} and Σ_{CE} - covariance matrices of equities and currencies,

Σ_{CC} - covariance matrix of the currencies only.

Importantly, if we consider the optimally currency hedged positions of the equity indexes, we can run a regression of the underlying assets on those hedges (Khoury & Jorion, 1996, 287). The slope coefficient of that regression is defined as β and Σ_{EC} as the covariance matrix of the underlying asset returns conditional on the hedges. Importantly, the beta coefficient corresponds to the position in forward contracts that provide the minimum variance hedge against currency risk, and it can be estimated as:

$$\beta = \Sigma_{CC}^{-1}\Sigma_{EC}$$

$$\Sigma_{EC} = \Sigma_{EE} - \beta'\Sigma_{CC}\beta$$

This a result of the fact that the variance of the portfolio with a fixed positions in stocks is defined as $\sigma^2_{Portfolio} = w_E'\Sigma_{EE}w_E + w_C'\Sigma_{CC}w_C - 2w_C'\Sigma_{CE}w_C$ and then minimum $\sigma^2_{Portfolio}$ is obtained when $2\Sigma_{CC}w_C - 2\Sigma_{CE}w_E = 0$ or when $w_E = \Sigma_{CC}^{-1}\Sigma_{CE}w_C = \beta w_E$ (Khoury & Jorion, 1996).

The optimal portfolio positions can be therefore extracted from the inverse of the abovementioned, partitioned covariance Σ matrix:

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{EC}^{-1} & -\Sigma_{EC}^{-1}\beta' \\ -\beta\Sigma_{EC}^{-1} & \Sigma_{CC}^{-1} - \beta\Sigma_{EC}^{-1}\beta' \end{bmatrix}$$

Then by simply rewriting we can obtain the formulas for the optimal equity and currency positions:

$$\begin{bmatrix} w_E = \Sigma_{EC}^{-1}\mu_E - \Sigma_{EC}^{-1}\beta'\mu_C \\ w_C = \Sigma_{CC}^{-1}\mu_C - \beta w_E \end{bmatrix}$$

This representation of the optimal minimum variance portfolio gives us the opportunity to consider the currencies positions individually and evaluate their two components separately. The first element, $\Sigma_{CC}^{-1}\mu_C$ is the speculative component which is present in the case when currency returns are different from zero. Interestingly, Glen & Jorion (1993) evaluate both the hedging and the speculative positions and find that the improvements in the risk-adjusted returns are mainly due to the hedging component. However, their results are likely to be significantly biased as they test their model in-sample.

In this paper, our focus is on the second element βw_E which is the hedging component, reducing the variance of our internationally diversified portfolio. By choosing to ignore the speculative demand, we implicitly make the assumption that the currency risk is not priced in the financial markets. We see that as plausible since the focus of this paper is the currency hedging and the risk management perspective of the portfolio. In other words, we consider the currencies as purely hedging instruments and instead we focus on the improvement of the optimization and the reduction of the total portfolio risk. Looking at the hedging demand independently should help us better understand and evaluate the risk reduction effect of the portfolio by looking also at the distribution characteristics and the implications of the applied methods on those.

2.2. Optimal hedges under different assumptions

The implementation of the currency hedge is a subject of dispute among practitioners and academia. The optimality condition depends largely on the assumptions, the assets correlation, and the goals of the investors.

If we assume no correlation, the optimal currency hedging is the simple unitary hedge $w_C = -w_E$, which will remove the variance of the currency from the portfolio returns (Solnik, 1974). Academic supporters of the unitary hedge argue that the long-term currency returns are zero and thus, they are contributing to the reduction in volatility without affecting returns. For example, Perold & Schulman (1988) show empirically that US investors can reduce the risk by applying full unitary hedges. At the same time, they argue that the premiums for currency holders are not far from zero and can be explained by time-varying premiums rather than stationary premiums. However, later studies, such as De Roon et al. (2012) argue that currency hedging is no free lunch since there is increasing evidence for currency premium and in the presence of such, currency risk hedging may affect the expected return on the portfolio in addition to its volatility.

Furthermore, the portfolio asset class composition and exposure also play a role in the choice of the hedge. For example, Campbell et al. (2010) find that in bond portfolios, the optimal hedge is close to a unitary hedge due to the low correlation between bond returns and currency returns. On the other hand, they also find that the minimum variance hedging portfolio differs from a unitary hedge, particularly for equities and that it significantly reduces the portfolio volatility. This is due to the presence of a significant correlation (both positive and negative) between equities and currencies. However, it is worth mentioning that their analysis is based on in-sample data. As they test their model on the data that was used to estimate it and this guarantees that the hedged portfolio will display lower volatility than the unhedged portfolio. Out-of-sample testing mimics a realistic implementation of the hedge and provides a better assessment of its economic benefits.

Moreover, there is increasing empirical evidence suggesting that the currency risk is priced in the international markets (Dumas & Solnik, 1995). According to De Santis & Gérard (1998), the currency risk premium is large, economically significant, and varies over time. If currency risk is priced in the securities, then hedging will also affect the return of the portfolio in addition to the volatility. This will also mean that fully hedging the entire currency risk of a portfolio can also have a negative effect on portfolio returns since it fails to account for the expected returns of the currency exposure. Bearing some degree of exchange rate risk may actually improve the performance of international portfolios (De Roon et al., 2012). In addition, if currency risk is varying over time this implies that investors may be more or less exposed than their individual risk preferences, and therefore when constructing their optimal portfolio, they need to take into account their risk exposure when rebalancing their portfolios.

When considering a portfolio of foreign currency and the corresponding equities whose returns are correlated to varying degrees, the optimal weights become the weights of the minimum variance hedged portfolio. Therefore, the optimal hedge becomes the minimum variance hedge (MVH) which we choose to present as gamma and involves an optimization problem (De Roon et al., 2012, p.5).

$$w_{hedge} = -\beta w_E = -\Gamma\, w_E = -\Sigma_{CC}^{-1}\Sigma_{EC}w_E \tag{1}$$

$$\Gamma = \Sigma_{CC}^{-1}\Sigma_{EC}$$

We can see that the hedge depends on the covariance matrix between the currencies and equities and the inverse of the covariance matrix of the currencies. As mentioned above, the optimal hedge is also the beta in the OLS regression of the unhedged portfolio returns on a constant and the currency returns (De Roon et al., 2012). This means that when the unhedged equity returns are positively correlated with the corresponding currency returns, the currency exposure is hedged by imposing a negative weight in the hedging portfolio.

The covariance matrix, which we can see is the key to the optimal currency hedge, is prone to estimation errors, and those are further amplified by the estimation of its inverse and the increase of the number of assets included in the portfolio. Since financial data is usually noisy, heavily skewed, and characterized by fat tails, estimators tend to contain significant amounts of noise. This noise creates problems in portfolio construction, resulting in estimation errors and poor portfolio performance. The problem becomes particularly severe for estimation involving a large number of parameters and low numbers of historical observations, which is usually the case in portfolio construction.

For example, DeMiguel et al. (2009) evaluate the performance of the sample-based mean-variance model against naive diversification. In their evaluation of 14 models across seven empirical datasets, they find that the gain from optimal diversification is more than offset by estimation errors. The paper discusses that the reason for the poor performance is largely due to the high sensitivity of the means and covariance estimates, where small changes to these inputs result in extremely different portfolios.

Furthermore, in another paper named “The Markowitz Optimization Enigma: Is 'Optimized' Optimal?” Michaud (1989) describes the mean-variance portfolio optimization as error maximization since the framework tends to maximize the effects of errors in the input assumptions. Nevertheless, Michaud underscores that the framework is superior to many ad hoc techniques and its performance may be enhanced by the sophisticated adjustment of inputs and the imposition of constraints based on fundamental investment considerations and the importance of priors.

The research suggests that reduction of the noise and adjustments of the inputs may lead to improvement of the financial estimators and therefore to more stable and better-performing portfolios, both with respect to risk and returns.

As we saw above, the optimization problem of the minimum variance currency hedge for an internationally diversified portfolio is multidimensional and involves estimation of the covariance matrix of the assets and the currencies and its inverse. In order to deal with the problems arising in high dimensional regimes, statisticians often apply the so-called sparsity or regularization methods which impose a penalty on the estimated parameters to control the magnitude of those. One of the first to propose such a method is Dempster (1972) as he suggests reducing the elements of the inverse covariance matrix to zero. Subsequently, numerous approaches for sparse matrix estimators have been developed and proposed and some of those have been found to significantly improve the estimates and reduce the portfolio risk (Goto & Xu, 2015; Millington & Niranjan, 2017).

We hypothesize that the application of penalization in the hedging element of the portfolio will lead to a reduction of the estimation errors and improvement of the portfolio risk. We apply and test the performance of several methods for the penalization of the covariance matrix and its inverse including, Ledoit & Wolf (2004) penalization, as well as the graphical lasso (glasso) estimator, proposed and evaluated by Friedman et al. (2008) and applied by Goto & Xu (2015). By including shrinkage in the estimation, we aim to minimize the impact of the data noise and multicollinearity to reduce the volatility of our portfolio. We discuss and compare different maximization techniques and evaluate penalization parameters to solve the problem of finding the best estimator.

As mentioned above, some regularized approaches achieve significant out-of-sample risk reduction and improve certainty equivalent returns after transaction costs when applied to an entire portfolio. For instance, Goto & Xu (2015) argues that gains in the portfolio risk reduction from glasso estimator are particularly strong for portfolios where the hedging relations tend to be more difficult to estimate (as is the case of portfolios of individual stocks). Moreover Millington & Niranjan, (2017) argue that the portfolios constructed using the graphical lasso have a lower variance of risks and returns than those constructed using the empirical covariance, particularly when applied out-of-sample.

However, to our knowledge, those approaches have not been tested in portfolio construction involving a hedging element. Our aim is therefore to apply the penalization methods, particularly with the aim to minimize the noise effect in the optimal currency hedge and to decrease the risk of the portfolio. We compare the risk reduction of our internationally diversified equally weighted portfolio, hedged with the minimum variance hedge, and then apply the penalization on it with the aim to reduce the estimation error. For comparison, we also evaluate the performance of the unitary hedge and Black´s universal hedge.

3. Data

We focus on developed countries with floating exchange rates, mostly driven by supply and demand and with little Central bank interventions. Our dataset consists of monthly data for Norway, as a domestic investment, and the G10 countries' local indexes as our foreign investment and the corresponding foreign currency exchange rates. The list can be found in Appendix Table 1. As we are considering the investments from the perspective of a Norwegian investor, we estimate the returns of the foreign equities with the following formulas:

$$R_{E,t+1} = \left(\frac{P_{t+1} + D_{t+1}}{P_t} - I_{t+1}^{Country}\right)\left(\frac{X_{t+1}}{X_t}\right) - 1$$

$$R_{C,t+1} = \frac{X_{t+1}\left(1 + I_{t+1}^{Country}\right)}{X_t} - \left(1 + I_{t+1}^{Norwegian}\right)$$

Where:

R_E – is the equity index return at the corresponding time-period,

R_C – is the currency return at the corresponding time-period,

X_t and X_{t+1} - is the exchange rate between foreign currency and Norwegian krone at the corresponding time-period,

P_t and P_{t+1} – is the price of the index in the local currency at the corresponding time-period,

D_t and D_{t+1} – is the index's dividends at the corresponding time-period,

$I_{t+1}^{Country}$ – is the local risk-free interest rate at the corresponding time-period,

$I_{t+1}^{Norwegian}$ – is the domestic risk-free interest rate at the corresponding time-period.

The data for the equity indexes is retrieved from Refinitiv Eikon and we chose to use the corresponding MSCI Country Price Index for each individual country since using the same provider for the indexes makes those comparable with each other. For the currencies we use Norges Bank exchange rates database where we use end of month values to match the data frequency from Eikon. The Norwegian krone is our domestic, quote currency, and the other currencies are base currencies since they represent the number of NOK the Norwegian investor needs to buy 1 unit of the foreign currency. For the risk-free interest rates, we use 1-month interbank deposit rates for the corresponding currency. We also consider the interest rate differentials to remove the effect of high vs. low-interest rates countries. The fact that domestic currency appreciates when domestic nominal interest rates exceed foreign interest rates, is also known as the forward premium puzzle or the Fama puzzle in currency markets. We used nominal returns as taxes, investment fees, and inflation is not factored into those. Hence, they give a better indication of how the investment will perform in the future and not in the past.

Our sample covers the period between February 1986 and April 2022, a total of 435 monthly observations. We choose to use monthly frequency in order to have as little data noise as possible. The daily and weekly observations are much more noisy and full of random shocks than the monthly data. Other papers focused on hedging use also monthly, and some even prefer quarterly observations. However, as we also want to have a good number of observations, we find the monthly frequency optimal. Since our analysis is focused on currency hedging, we assume that hedging transactions take place once a month.

We divide our sample into training and testing periods with a 70/30 proportion. For the initial covariance matrix estimation, we use the data from February 1986 till February 2011. Moreover, for out-of-sample testing of the optimizer, we use the data period from March 2011 to April 2022. We rebalance our out-of-sample portfolio every month by adding the past month to our training data set in a recursive manner and we reestimate the weights of the portfolio for the following month.

3.1. Descriptive Statistics

The descriptive statistics for the entire dataset are presented in Table 3.1. The average monthly returns from the perspective of a Norwegian investor for the Swiss franc, the Japanese yen, the Canadian dollar, the Swedish krona, the Euro, and the US dollar are all negative and close to 0. The returns for the British pound are positive but still insignificantly small. However, one could argue that the median is the most proper measure for estimating expected returns in financial data as it is not influenced by extreme observations. Looking at the medians of the currency distribution we can see that all those except for the Canadian dollar are negative and close to zero.

Asset	Mean	St. dev.	Median	Min	Max
MSCI Norway Price Index	0,24%	6,24%	0,77%	-31,01%	15,47%
MSCI Canada Price Index	0,34%	4,75%	0,75%	-22,28%	14,95%
MSCI France Price Index	0,54%	5,77%	1,18%	-22,50%	23,40%
MSCI Germany Price Index	0,41%	6,35%	0,80%	-26,54%	18,27%
MSCI Japan Price Index	0,23%	5,91%	0,28%	-21,33%	19,82%
MSCI Netherlands Price Index	0,58%	5,58%	0,95%	-23,14%	16,46%
MSCI Italy Price Index	0,29%	6,71%	0,27%	-16,69%	28,13%
MSCI Belgium Price Index	0,40%	5,74%	0,79%	-23,79%	27,06%
MSCI Switzerland Price Index	0,66%	4,60%	0,87%	-23,57%	16,70%
MSCI UK Price Index	0,16%	4,29%	0,24%	-25,80%	14,78%
MSCI USA Price Index	0,55%	4,67%	0,96%	-22,30%	15,41%
MSCI Sweden Price Index	0,63%	6,52%	0,90%	-22,77%	28,70%
CHF	-0,04%	1,98%	-0,20%	-7,74%	11,86%
JPY	-0,29%	2,01%	-0,36%	-7,94%	6,41%
CAD	-0,02%	2,20%	0,00%	-5,17%	7,30%
EUR	-0,03%	2,05%	-0,19%	-5,95%	9,94%
GBP	0,07%	1,96%	-0,01%	-10,54%	12,12%
SEK	-0,10%	1,60%	-0,10%	-10,23%	8,10%
USD	-0,09%	2,54%	-0,12%	-6,15%	13,52%

Table 3.1. Descriptive statistics

3.2. Correlations

The correlations are undisputedly more interesting to look at when we discuss hedging. Table 3.2. report the correlation matrix estimated over the entire sample. We can confirm the lower correlation across countries confirming the discussion about the benefits of international diversification. Since the idea of diversification is mathematically based on the correlation between the holdings, investors aiming to reduce the risk of their portfolios should aim to combine assets that are low to negatively correlated. As Roll (1992) points out, the intercorrelation among markets is surprisingly low given the global financial integration. To some degree, we can say that European markets are more correlated with each other, which is also another reason to globally spread our Norwegian portfolio.

With respect to the individual currencies, all indexes are positively correlated with their corresponding local currencies. This means that when the indexes are increasing, the currencies are as well and vice versa. This implies that without hedging in bull markets the returns of the portfolio from the domestic perspective will be higher and in bear markets they will be lower. Therefore, due to the positive correlations, we expect to obtain mostly negative weights in the considered below hedging portfolios, implying that most of the currencies must be short-selled. However, in a portfolio consisting of multiple assets the cross correlations are very important for deciding the correct hedging position and therefore we cannot draw direct conclusions by simply looking at the correlation matrix.

	MSCI Norway Price Index	MSCI Canada Price Index	MSCI France Price Index	MSCI Germany Price Index	MSCI Japan Price Index	MSCI Netherlands Price Index	MSCI Italy Price Index	MSCI Belgium Price Index	MSCI Switzerland Price Index	MSCI UK Price Index	MSCI USA Price Index	MSCI Sweden Price Index	CHF	JPY	CAD	EUR	GBP	SEK	USD
MSCI Norway Price Index	1,00																		
MSCI Canada Price Index	0,64	1,00																	
MSCI France Price Index	0,60	0,66	1,00																
MSCI Germany Price Index	0,60	0,66	0,86	1,00															
MSCI Japan Price Index	0,42	0,45	0,48	0,43	1,00														
MSCI Netherlands Price Index	0,63	0,69	0,82	0,84	0,46	1,00													
MSCI Italy Price Index	0,51	0,54	0,76	0,71	0,45	0,69	1,00												
MSCI Belgium Price Index	0,59	0,59	0,77	0,73	0,41	0,79	0,64	1,00											
MSCI Switzerland Price Index	0,51	0,53	0,67	0,66	0,45	0,70	0,57	0,61	1,00										
MSCI UK Price Index	0,57	0,54	0,63	0,59	0,39	0,63	0,52	0,56	0,65	1,00									
MSCI USA Price Index	0,55	0,75	0,73	0,73	0,45	0,79	0,56	0,68	0,67	0,64	1,00								
MSCI Sweden Price Index	0,60	0,61	0,68	0,71	0,49	0,69	0,60	0,58	0,58	0,57	0,64	1,00							
CHF	-0,22	-0,12	-0,07	-0,08	-0,02	-0,04	-0,07	-0,09	0,24	0,04	0,03	-0,07	1,00						
JPY	-0,03	0,12	0,12	0,15	0,34	0,19	0,08	0,09	0,21	0,02	0,20	0,11	0,42	1,00					
CAD	0,07	0,45	0,31	0,36	0,15	0,41	0,24	0,30	0,19	0,08	0,40	0,25	0,16	0,41	1,00				
EUR	0,00	0,22	0,31	0,34	0,06	0,42	0,21	0,28	0,25	0,08	0,35	0,12	0,43	0,50	0,72	1,00			
GBP	-0,26	-0,19	-0,14	-0,11	-0,04	-0,13	-0,13	-0,18	0,05	0,18	-0,10	-0,09	0,53	0,35	0,16	0,29	1,00		
SEK	0,02	0,16	0,10	0,13	0,12	0,17	0,10	0,12	0,24	0,14	0,20	0,23	0,43	0,38	0,33	0,36	0,41	1,00	
USD	-0,09	0,18	0,25	0,28	0,04	0,35	0,17	0,22	0,20	0,02	0,36	0,11	0,37	0,50	0,77	0,87	0,26	0,31	1,00

Table 3.2 Correlation matrix

4. Methodology

4.1. Important characteristics of the covariance and concentration matrices

The covariance matrix and its inverse are heavily used in portfolio construction and are central for portfolio risk minimization. Since the minimum variance portfolio does not consider returns but rather relies only on the covariance matrix estimation. The inverse covariance matrix, also known as the concentration matrix, contains information about the partial correlation between variables or put differently it shows how correlated are two assets given the presence of other assets in the portfolio. According to Stevens (1998) it determines the optimal holding of a given risky asset, the slope of the risk-return efficiency, and the pricing of risky assets in the Capital Asset Pricing Model (CAPM). Stevens also points out that the i-th row (or column) of the concentration matrix is proportional to the stock's minimum variance hedge portfolio. The hedge portfolio consists of a long position in the i-th stock, and a short position in the "tracking portfolio" of the other stocks to hedge the i-th stock.

While diagonal elements of the covariance matrix are the individual asset variances (the covariance of each asset with itself), the off-diagonal elements are the covariances between those. Importantly, higher than 0 off-diagonal elements, or in other words, covariances between assets higher than 0, are resulting in extreme values when inverting the covariance matrix. Those extreme values could produce unstable and often unreasonable portfolio weights and result in poor portfolio performance and higher risk.

4.2. Hedging techniques

We consider several hedging approaches and apply it to our internationally diversified portfolio to hedge it against currency risk:

1) Unitary hedge

When the unitary hedge approach is applied, the full-face value of the foreign investment is hedged in the forward market on one-to-one basis (Khoury & Jorion, 1996, 298). This hedging technique does not consider the interdependencies between the assets in the portfolio and thus, can lead to over hedging and poor performance when correlations between currencies and equities are high. Hence, the weights of the hedging portfolio are defined as following:

$$w_C = -w_E$$

2) Black´s universal hedge

We also apply the universal hedge which was first proposed by Fischer Black. In his paper Black (1989) argues that, given no barriers to international investment, all the investors want to hold a share of a fully diversified world equity portfolio. Since some investors must lend when others borrow, in equilibrium they both need to hedge equally, in proportion to their stock holdings. Based on these assumptions, he derived the formula for the universal hedge:

$$\frac{\mu_e - \sigma_e^2}{\mu_e - 0.5\sigma_{er}^2}$$

Where:

μ_e - the average of the expected excess returns on the equity portfolio

σ_e - the average volatility of the equity portfolio

σ_{er} - the average exchange rate volatility across all pairs of countries

Hence, the hedge is applied as:

$$w_C = -\frac{\mu_e - \sigma_e^2}{\mu_e - 0.5\sigma_{er}^2} w_E$$

In other words, the Black's universal hedge concept is very similar to the unitary hedge, however, the index exposure hedge ratio is not 100%, but the optimal one, which could be estimated with the formula above.

3) Minimum variance hedge (MVH)

In this paper we give most attention to the minimum variance hedge, as is the only one that takes into account the interdependencies in the portfolio. We focus only on the hedging portfolio part, and we assume zero currency returns, hence the formula of equation (1), presented in the literature review, is used to estimate the currency portfolio weights:

$$w_C = -\Sigma_{CC}^{-1}\Sigma_{CE}w_E$$

Where:

w_C - is the minimum variance optimal position in currencies,

Σ_{CC}^{-1} - is the currency concentration matrix,

Σ_{CE} - is the covariance matrix between indexes and currencies,

w_E - are the weights of the indexes' portfolio.

We intend to improve the MVH by applying shrinkage techniques to the currency concentration matrix Σ_{CC}^{-1} as well as to the entire covariance matrix.

4.3. The estimators for the MVH

As was mentioned before the practice to estimate the covariance matrix on historical data makes the estimates uncertain and prone to estimation errors. Furthermore, since the equity markets are correlated to varying degrees, the data is characterized by multicollinearity. Multicollinearity means also that the off-diagonal elements of the covariance matrix and its inverse are susceptible to large estimation errors (Goto & Xu, 2015). Shrinkage methods are often applied to approach the issue of such errors and are particularly useful in the mitigation of extreme values resulting in the inversion of the covariance matrix described above.

There is an ongoing discussion if the shrinkage is best applied to the covariance matrix or its inverse. As Goto & Xu (2015) point out, the relative performance of the portfolios may depend on datasets, estimation windows, and performance measures as well as testing methodologies. This means that the performance of the different methodologies may be different across different data sets and time frame windows. Moreover, Bien & Tibshirani (2011) state that while the zeros in the inverse covariance matrix correspond to conditional independencies between variables, zeros in the covariance matrix correspond to marginal interdependencies between variables.

Therefore, we choose to test both approaches by applying shrinkage of the inverse currency partition of the covariance matrix which is used in the optimal hedge using the approach of Goto & Xu (2015), Yuan & Lin (2007), and Friedman et al., (2008), as well as the approach of Ledoit & Wolf (2004) for shrinkage of the entire covariance matrix, which is then partitioned and inverted. The aim of both is to achieve an estimator which is sparse, or with other words, an estimator whose less important off-diagonal elements will be shrunk towards zero.

4.3.1. Shrinkage on the inverse covariance matrix

The methodology we apply to shrink the concentration matrix has been used in several other papers such as Friedman et al., (2008), Goto & Xu (2015), Millington & Niranjan (2017), and others, however, we are the first in our knowledge to apply it in the context of currency risk hedging.

Given the observations of the currency portfolio returns, which are i.i.d and assumed to come from N-dimensional multivariate normal distribution, we aim to estimate the inverse covariance matrix for currencies by quasi maximum likelihood (QML) estimator. Particularly, instead of applying lasso to each hedge regression to estimate each row/column of the concentration matrix' estimator, we want it to be estimated jointly to restrict the estimator to be positive, definite, and symmetric. This approach is expected to result in a portfolio with more stable weights and therefore - significantly improve portfolio performance.

Similar to Friedman et al. (2008) and Millington & Niranjan (2017), we use log-likelihood with a penalty $\rho \geq 0$ on the overall size of the hedge trades to estimate all elements of the concentration matrix estimator Ψ with the following formula:

$$\max_{\Psi} \ln \det(\Psi) - trace\left(\hat{\Sigma}\Psi\right) - \rho \sum_{i=1}^{N} \sum_{j=1, i \neq j}^{N} |\psi_{ij}| \qquad (3)$$

In their paper Goto & Xu (2015) use a slightly different formula to estimate the concentration matrix:

$$\max_{\Psi} \frac{T}{2} \ln \det(\Psi) - \frac{T}{2} trace(\hat{\Sigma}\Psi) - \rho \sum_{i=1}^{N} \sum_{j=1, i\neq j}^{N} |\psi_{ij}| \tag{4}$$

For both formulas:

$\hat{\Sigma}$ - the estimated sample covariance matrix

Ψ - the estimator of the concentration matrix

N – number of assets,

T – number of the data points.

The only difference between the two approaches is the way they are discarding the constant. For simplicity, the following discussion will be based on formula (3). Moreover, the degree of the penalization is driven by the penalization factor ρ. Larger values of ρ promote more sparsity and shrinkage of the estimator, whereas zero penalization makes the solution identical to the unpenalized concentration matrix. There are different ways to evaluate the optimal penalization parameter. We choose to use the grid search and Bayesian information criterion (BIC) criteria to search for the value that optimizes the predictive likelihood for our training period. According to Goto & Xu (2015) the estimator can deliver consistent performance when the optimal value of ρ remains stable over time.

There are several ways to find the solution to the QML estimation problem. In this paper, we will consider two of them, specifically the lasso (glasso) algorithm proposed by Friedman et al. (2008) as well as the application of standard maximization techniques.

Friedman et al. (2008) state that the QML estimation problem looks like N-coupled Lasso least-squares problem. They use the block-wise coordinate descent approach proposed by Banerjee et al. (2008) as a starting point and propose a new recursive algorithm for this problem. The proposed of them glasso sweeps across each row, solving the individual lasso problems while keeping the others fixed, and then repeats until all the coefficients converge. Importantly, the algorithm computes the estimator for the inverse covariance matrix after convergence, as it works on the estimated sample covariance matrix $\hat{\Sigma}$ and then recover relatively cheaply the estimator Ψ (Friedman et al., 2008).

We also apply standard maximization techniques to the QML problem to test their performance. When it comes to those, there are two main types - derivative-based and stochastic-based maximizations. In both cases, the large number of parameters significantly complicates the problem since the function can be multimodal. The optimization is particularly problematic in the case of the derivative-based methods when the maximization search may stop at a local maximum point instead of the global max. The problem can be mitigated by choosing a sensible starting point and the appropriate algorithm depending on the problem. As an initial starting point for the QML maximization problem, we will use our sample concentration matrix. To evaluate the robustness of the estimator, we applied several maximization algorithms which are presented in Table 4.1.

Method	Description
Sequential quadratic programming (SQP)	Deterministic, uses derivatives. If the problem is unconstrained, then the method reduces to Newton–Raphson method.
Active-set (AS)	Deterministic, uses derivatives.
Nelder-Mead	Deterministic, derivative-free.

Table 4.1. Tested maximization methods

The SQP and AS performances are similar, but the SQP proves to perform slightly better. The Nelder-Mead method performs poorly, and we discard it. Therefore, we chose to base our analysis on the SQP maximization method, which is regarded as one of the most successful methods for the numerical solution of constrained nonlinear optimization problems.

The above estimators (glasso and SQP) will be used instead of the sample concentration matrix Σ_{CC}^{-1} to estimate the currency hedge weights. However, we are unable to penalize the entire hedge using the glasso estimator due to the structure of the minimum variance hedge. The application of glasso algorithms requires a NxN square matrix in order to estimate the determinant of the matrix as a part of the equation (3). As described above, the minimum variance hedge is estimated as:

$$\Gamma = \Sigma_{CC}^{-1}\Sigma_{CE}$$

While our inverse matrix of the currencies Σ_{CC}^{-1} is square (7x7), the matrix of the currencies and equities Σ_{CE} is not (in our case it is 13x7), making it impossible to apply the optimization algorithm on the entire hedge. We provide further discussion of the problem in the results section.

4.3.2. Shrinkage on the covariance matrix

The methodology we apply to shrink the covariance matrix was first proposed by Olivier Ledoit in his Ph.D. thesis Essays on Risk and Return in the Stock Market (1995). Later it was further discussed by him and Michael Wolf in their paper A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices (Ledoit & Wolf, 2004). To get a well-conditioned structured estimator they assumed that all variances are the same and all covariances are zero. Hence, the shrinkage target is to keep the covariance matrix diagonal, and shrinkage intensity influences only the off-diagonal elements. In other words, we again shrink only the off-diagonal elements of the covariance matrix. After the application of the shrinkage, we partition the covariance matrix to extract the elements Σ_{CC} and Σ_{CE} needed for the minimum variance hedge. We then take the inverse of the already shrank Σ_{CC} and apply it to estimate the hedge. This means that in this approach both elements of the hedge are equally shrunk.

5. Results and Analysis

We choose to present the applied estimators' results separately for our hedging portfolio and for our total portfolio in order to evaluate their performance both on an individual basis and on the total portfolio risk profile. The presented results are estimated on our out-of-sample monthly data for the period from March 2011 till April 2022. The results of the out-of-sample data are calculated in a recursive manner by monthly rebalancing and gradual expansion of the training period. As mentioned above, we used grid search and Bayesian information criterion (BIC) to estimate the value of the optimal penalization parameter. The optimal values for ρ which we obtained are 0.1 the SQP method and 0.0001 for the glasso algorithm.

We evaluate the risk reduction by measuring three dimensions of the out-of-sample portfolio risk profile: the volatility, the kurtosis, and the skewness. Including higher moments of the portfolio, distribution to evaluate the risk is important since the reduction of volatility itself does not guarantee the reduction of the total portfolio risk. A portfolio may have lower volatility but also higher kurtosis which in combination with negative skew, implies that the tail risk of the portfolio is higher indicating that a "black swan" event has a higher chance of occurring.

5.1. Currency Hedge results

Looking exclusively at our currency minimum variance optimal hedge portfolio (MVOH), presented in Table 5.1., we can see that we obtain positive returns, skewness close to zero, and low kurtosis. In fact, as an individual portfolio, it has the highest return and SR of all the applied techniques, proving it difficult to improve further. Moreover, all the other applied hedging adjustments deliver negative returns except for the Ledoit-Wolf (LW) estimator. The worst performing is the QML SQP maximization, which increases the standard deviation by threefold.

In terms of standard deviation, surprisingly, the Black´s Universal hedge proves to be the lowest. However, this comes at the cost of a significant increase in tail risk and negative skewness, meaning that the hedge portfolio is more susceptible to market crashes. The unitary hedge performs similarly with a slightly higher standard deviation and similar distribution characteristics.

While the minimum variance hedge proves its optimality on an individual basis, we need to look at the overall portfolio to draw further conclusions.

Currency hedge	Optimal Hedge	Unitary Hedge	Universal Hedge	SQP	Glasso	Ledoit-Wolf
Return	0,07%	-0,14%	-0,11%	-0,38%	-0,06%	0,06%
Standard deviation	1,70%	1,55%	1,25%	5,50%	2,01%	1,64%
Sharpe ratio	4,06%	-8,85%	-8,85%	-6,90%	-3,20%	3,59%
Skewness	-0,05	-0,86	-0,86	-0,37	-0,17	-0,06
Kurtosis	3,54	5,77	5,77	3,58	3,74	3,52

Table 5.1. Currency hedge portfolio results

5.2. Total portfolio results

Examining the total portfolio results, reported in Table 5.2., we can see that the minimum variance hedge delivers the results it promised, by both increasing the returns and decreasing the volatility of our currency-hedged portfolio. However, it also contributes to a slightly higher kurtosis, which means that it increases the tail risk making the portfolio more susceptible to sudden market crashes.

Portfolio	Unhedged Equity	Optimal Hedge	Unitary Hedge	Universal Hedge	SQP	Glasso	Ledoit-Wolf
Return	0,7%	0,7%	0,5%	0,6%	0,3%	0,6%	0,7%
Standard deviation	3,5%	3,3%	3,7%	3,6%	5,8%	3,5%	3,3%
Sharpe ratio	19,2%	23,0%	14,6%	15,7%	5,2%	17,4%	22,7%
Skewness	-0,10	-0,05	-0,56	-0,46	-0,73	-0,42	-0,09
Kurtosis	3,41	3,61	5,32	4,89	5,53	4,51	3,64

Table 5.2. Portfolio results

Contrary to our initial hypothesis, we do not find any significant improvements in the risk characteristics of the total portfolio by applying penalization techniques to our hedging element. Furthermore, the unitary and the universal hedges prove to be suboptimal when applied to equity portfolios as they not only decrease the expected returns but also significantly increase the standard deviation. As mentioned above, the unitary and the universal hedge are quite similar. We can see that the universal hedge performs slightly better than the unitary due to its structure since the application of the optimal hedge ratio in a way decreases the over hedging of the unitary hedge. However, the results are still worse than with the optimal hedge.

Moreover, we can see that by applying the unitary and universal hedges to our portfolio, the tail risk is increased drastically, and the returns distribution becomes more negatively skewed. This can be explained by the presence of significant correlation between both the equity indexes and their corresponding currencies, which neither of the two hedging techniques take into account. For instance, Glen & Jorion (1993) points out, the unitary hedge ratio is suboptimal since it ignores correlations between exchange rates and local returns as well as speculative motives for taking currency positions. Solnik (1974) argues similarly, that a unitary hedge is optimal only when currencies and equities are uncorrelated, and the risk management demands for foreign currencies are zero.

With respect to the SQP maximization, the results are even more dramatically aggravated although we applied the grid search method to estimate the best penalization parameter. In fact, the SQP presents the worst results with respect to risk increase by doubling the risk, heavily negatively skewing the returns, and significantly increasing the tail risk of our portfolio. The SQP maximization performance can be explained due to the nature of the problem that is intended to solve. As Bien & Tibshirani (2011) discuss, the minimization of equation (3) is a formidable challenge since the objective function is non-convex and therefore may have many local minima. This is also their main motivation for developing the glasso algorithm.

Of all the applied penalization methods, Ledoit-Wolf is the only which delivers some improvement to the risk characteristics of our portfolio, although they are statistically and economically insignificant. The decrease of the volatility is coming at the cost of the returns, resulting in the same Sharpe Ratio as with the optimal hedge achieving therefore a zero risk-adjusted benefit for the investor. As mentioned above, the method is also the only method that is applied to the entire sample covariance matrix where the matrix is penalized, before being partitioned and used in the estimation for the hedge, resulting in the entirely penalized hedge. We could argue that the penalization of the entire hedge is therefore preferable, however, as the results are not significant, this will be a speculation.

In order to evaluate the reasons for the rather surprising results we decided to present the covariance and the concentration matrices used in the construction of the hedge, the statistics for the determinant of the estimator ψ_{CC} and the trace of the $\widehat{\Sigma_{CC}}\psi_{CC}$, as well as the currency weights at the end of our testing period (30.04.22).

Inverse currency sample covariance matrix	**CHF**	**JPY**	**CAD**	**EUR**	**GBP**	**SEK**	**USD**
CHF	4 683	- 691	1 488	- 1 484	- 1 550	- 1 116	- 476
JPY	- 691	3 852	- 282	- 273	- 421	- 666	- 735
CAD	1 488	- 282	5 835	- 1 519	143	- 1 165	- 2 940
EUR	- 1 484	- 273	- 1 519	10 854	24	- 400	- 6 013
GBP	- 1 550	- 421	143	24	3 940	- 953	- 110
SEK	- 1 116	- 666	- 1 165	- 400	- 953	5 631	720
USD	- 476	- 735	- 2 940	- 6 013	- 110	720	8 044
Estimator SQP	**CHF**	**JPY**	**CAD**	**EUR**	**GBP**	**SEK**	**USD**
CHF	4 372	- 19	- 166	- 122	- 196	7	- 20
JPY	- 19	3 580	1	4	- 2	- 20	- 167
CAD	- 162	1	5 341	- 344	94	2	- 1 891
EUR	- 122	4	- 344	10 349	- 2	- 67	- 4 977
GBP	- 196	- 2	94	- 2	3 724	15	- 22
SEK	7	- 20	2	- 67	15	5 487	- 190
USD	- 21	- 167	- 1 891	- 4 977	- 21	- 190	7 265
Estimator glasso	**CHF**	**JPY**	**CAD**	**EUR**	**GBP**	**SEK**	**USD**
CHF	2 910	- 293	0	- 299	- 735	- 264	- 97
JPY	- 293	2 805	0	- 191	- 138	- 133	- 525
CAD	0	0	3 251	- 616	- 0	- 0	- 1 331
EUR	- 299	- 191	- 616	4 556	- 0	- 0	- 2 098
GBP	- 735	- 138	- 0	- 0	2 856	- 189	0
SEK	- 264	- 133	- 0	- 0	- 189	3 981	- 97
USD	- 97	- 525	- 1 331	- 2 098	0	- 97	3 524
Inverse currency covariance matrix Ledoit-Wolf	**CHF**	**JPY**	**CAD**	**EUR**	**GBP**	**SEK**	**USD**
CHF	4 575	- 676	1 379	- 1 384	- 1 505	- 1 078	- 459
JPY	- 676	3 809	- 287	- 301	- 417	- 652	- 693
CAD	1 379	- 287	5 572	- 1 523	141	- 1 087	- 2 702
EUR	- 1 384	- 301	- 1 523	9 977	4	- 358	- 5 334
GBP	- 1 505	- 417	141	4	3 892	- 935	- 101
SEK	- 1 078	- 652	- 1 087	- 358	- 935	5 565	622
USD	- 459	- 693	- 2 702	- 5 334	- 101	622	7 317

Table 5.3. Concentration matrices estimators ψ_{CC} as of 30.04.22

The inverse currency covariance matrix Σ_{CC} per 30.04.22 and its corresponding estimators, are presented in Table 5.3. We can see that the SQP minimization shrinks the concentration matrix of the currencies and as expected reduces the off-diagonal elements. However, it does not seem to find the optimal solution, leading to an increase in the risk of our currency hedge portfolio and of the corresponding total portfolio.

Furthermore, we can also observe that it changes the sign of some of the covariances. If the regressors (the currencies) are orthonormal, the lasso is shrinking them toward zero, but it never crosses zero or alternates the sign. However, as in today's global markets, we cannot talk about independence and uncorrelated indices, particularly not for country indices and their currencies, meaning that those are not orthonormal. The lack of orthogonality requires, therefore, the application of an iterative solution (Goto & Xu, 2013).

Such a solution is proposed by Friedman et al. (2008) whose algorithm proves to be remarkably fast and efficient. We can also notice that the performance of the glasso is superior to the simple minimization techniques. In Table 5.3 and 5.4 we observe that the approach shrinks efficiently the off-diagonal elements of the covariance matrix without affecting the diagonal and after convergence computes the inverse covariance matrix (Friedman et al., 2008). The glasso is however also exacerbating the results of the optimal hedge by increasing volatility, skewness, and tail risk for our overall portfolio.

As mentioned above, we also choose to present the currency covariance matrices $\widehat{\Sigma_{CC}}$ for the MOVH, the glasso and the LW as we find it useful to present and evaluate the effects of the shrinkage techniques on the covariance matrices, for those where shrinkage is applied to. This means that we do not present the covariance for the SQP as is the only method that works directly and exclusively on the inverse matrix Σ_{CC}^{-1}. In the Table 5.4 sample and shrinked covariance matrices we can clearly see the effect of the glasso shrinkage on the off-diagonal elements while the diagonal elements are preserved as in the original matrix, confirming that the algorithm is performing as expected. Importantly, the shrinkage for the glasso algorithm is significant while the off-diagonal elements are only very mildly shrinked by applying the LW approach.

Currency sample covariance matrix	CHF	JPY	CAD	EUR	GBP	SEK	USD
CHF	3,91E-04	1,69E-04	6,85E-05	1,74E-04	2,06E-04	1,35E-04	1,84E-04
JPY	1,69E-04	4,03E-04	1,81E-04	2,05E-04	1,39E-04	1,24E-04	2,57E-04
CAD	6,85E-05	1,81E-04	4,86E-04	3,25E-04	6,65E-05	1,15E-04	4,32E-04
EUR	1,74E-04	2,05E-04	3,25E-04	4,23E-04	1,17E-04	1,18E-04	4,55E-04
GBP	2,06E-04	1,39E-04	6,65E-05	1,17E-04	3,81E-04	1,27E-04	1,31E-04
SEK	1,35E-04	1,24E-04	1,15E-04	1,18E-04	1,27E-04	2,56E-04	1,28E-04
USD	1,84E-04	2,57E-04	4,32E-04	4,55E-04	1,31E-04	1,28E-04	6,47E-04

Currency covariance matrix glasso	CHF	JPY	CAD	EUR	GBP	SEK	USD
CHF	3,91E-04	6,86E-05	4,85E-05	7,39E-05	1,06E-04	3,53E-05	8,42E-05
JPY	6,86E-05	4,03E-04	8,41E-05	1,05E-04	3,87E-05	2,37E-05	1,57E-04
CAD	4,85E-05	8,41E-05	4,86E-04	2,25E-04	1,75E-05	1,49E-05	3,32E-04
EUR	7,39E-05	1,05E-04	2,25E-04	4,23E-04	2,53E-05	1,82E-05	3,55E-04
GBP	1,06E-04	3,87E-05	1,75E-05	2,53E-05	3,81E-04	2,72E-05	3,11E-05
SEK	3,53E-05	2,37E-05	1,49E-05	1,82E-05	2,72E-05	2,56E-04	2,80E-05
USD	8,42E-05	1,57E-04	3,32E-04	3,55E-04	3,11E-05	2,80E-05	6,47E-04

Currency covariance matrix Ledoit-Wolf	CHF	JPY	CAD	EUR	GBP	SEK	USD
CHF	3,91E-04	1,66E-04	6,75E-05	1,71E-04	2,03E-04	1,33E-04	1,81E-04
JPY	1,66E-04	4,03E-04	1,78E-04	2,02E-04	1,37E-04	1,22E-04	2,53E-04
CAD	6,75E-05	1,78E-04	4,86E-04	3,21E-04	6,55E-05	1,13E-04	4,26E-04
EUR	1,71E-04	2,02E-04	3,21E-04	4,23E-04	1,15E-04	1,16E-04	4,48E-04
GBP	2,03E-04	1,37E-04	6,55E-05	1,15E-04	3,81E-04	1,25E-04	1,29E-04
SEK	1,33E-04	1,22E-04	1,13E-04	1,16E-04	1,25E-04	2,56E-04	1,26E-04
USD	1,81E-04	2,53E-04	4,26E-04	4,48E-04	1,29E-04	1,26E-04	6,47E-04

Table 5.4. Covariance matrices estimators $\widehat{\Sigma_{CC}}$ as of 30.04.22

Moreover, we can notice an important change in the magnitude of the diagonal elements of the inverse covariance matrix in Table 5.3 which is caused by the shrinkage. This is due to the construction of the glasso algorithm, where the shrinkage is applied on the off-diagonal elements of the covariance matrix, while the diagonal elements remain unchanged. Consequently, the inverse is recovered through the algorithm (Friedman et al., 2008). However, this results in slightly different diagonal elements of the inverse and to changed order of magnitude of those. This is inevitable in the case of the glasso algorithm as it was developed by Friedman et al. (2008). The lasso SQP shrinkage seems to have lesser effect on the diagonal, however, it still affects the values and the magnitude for some of the currencies. This implies that shrinkage techniques change the order of magnitude and influence the interdependencies of the variables in the matrices used in the estimation of the minimum variance hedging portfolio in such a way that it leads to decrease of its efficiency. Importantly, for Ledoit-Wolf approach, as the only that does not exacerbate the results, the shrinkage is minimal, the diagonal is preserved and the order of magnitude in the inverse estimate is unchanged. Those results are again supporting the hypothesis that even small interference in the matrix is likely redundant and the more shrinkage is applied on the matrix, the worse the portfolio performance becomes.

We can also argue that the worsened results of the applied shrinkage methods may be due to the nature of the minimum variance hedge structure. The structures of both the QML problem and the glasso penalization algo make it appropriate only for problems involving NxN matrices. This means that one could only apply the method on our entire portfolio, as Goto & Xu (2013), or eventually on the entire hedging element of a portfolio with an equal number of currencies and equities since only then the covariance of those will result in a square matrix allowing penalization. Since the NxN currencies and equity portfolio is an unreasonable constraint and assumption, we choose to apply the penalization only on the inverted partition of the currencies Σ_{CC}^{-1} with the aim to minimize the noise in currencies data. However, we can see that the results prove to be suboptimal.

In order to further evaluate the results, we included the statistics for the determinant of the ψ_{CC} and the trace of the $\widehat{\Sigma_{CC}}\psi_{CC}$ presented in Table 5.5 and 5.6. Since shrinkage influences the values of those and both are part of the equation (3), the two elements can provide some insight on why the shrinkage does not improve the optimal hedge. In mathematics, both the determinant and the trace can indicate the characteristics of the changes in a given matrix. The statistics are estimated out-of-sample.

Hedge	Mean	St. dev.	Median	Min	Max
Optimal Hedge	1,94866E+25	1,63247E+24	1,91079E+25	1,71722E+25	5,73992E+25
SQP	7,34308E+25	9,71354E+24	7,1228E+25	5,899E+25	9,62964E+25
Glasso	2,08688E+24	2,24142E+23	2,04478E+24	1,74814E+24	2,60621E+24
Ledoit-Wolf	1,62066E+25	1,34903E+24	1,57474E+25	1,41374E+25	1,9614E+25

Table 5.5 Determinant of the ψ_{CC}

Hedge	Mean	St. dev.	Median	Min	Max
Optimal Hedge	7,00	0,00	7,00	7,00	7,00
SQP	10,96	0,11	10,94	10,69	11,23
Glasso	5,82	0,10	5,83	5,60	5,93
Ledoit-Wolf	7,00	0,00	7,00	7,00	7,00

Table 5.6. Trace of the $\widehat{\Sigma_{CC}}\psi_{CC}$

We can see that the determinant for the MVOH and the determinant of the LW optimized hedge are quite close to each other, with the LW being slightly lower. However, for the SQP and the glasso the determinants are driven in the opposite directions where the SQP has a drastic increase and the glasso a drastic decrease of the value of the determinant. We argue that the results suggest that shrinkage causes large deviations away from the optimality of the hedge, which leads to worse performance. This is further supported by the fact that the increase of the penalization leads to poorer performance.

The trace statistics for the $\widehat{\Sigma_{CC}}\psi_{CC}$ confirm further the similarity between the MVOH and the LW shrinkage. The means of those are equal and the standard deviation of those is essentially equal to zero. The reason for that can be explained by the fact that in LW the entire covariance matrix Σ is shrank and then partitioned. However, for the SQP method, the shrinkage is directly applied on the concentration matrix of the currencies Σ_{CC}^{-1}, while with the glasso the shrinkage is applied first on the covariance matrix of those Σ_{CC} and then the inverse Σ_{CC}^{-1} is recovered by the algorithm. This leads to different mean, median and volatility of the $\widehat{\Sigma_{CC}}\psi_{CC}$, supporting the former discussion about the bad performance being mainly due to disturbance in the minimum variance hedge structure.

Another way to explain the results is by looking at the simplified calculations involved in the estimation of the covariance matrix. We know that we can represent the volatility of the equity and currency portfolio with the following formula:

$$\sigma_{PTF} = \sqrt[2]{(w_E\sigma_E)^2 + \ 2w_E w_C cov_{E,C} + \ (w_C\sigma_C)^2}$$

Where:

σ_{PTF} – is the standard deviation of the constructed portfolio,

w_E – is the total weight of the equity index portfolio,
w_C – is the total weight of the currency portfolio,
σ_E and σ_C – are the corresponding volatilities of the portfolios
$cov_{E,C}$ – is the covariance between the portfolios.

While our equity portfolio is fixed, the currency portfolio' weights, standard deviation, and covariance with the indexes' portfolio depend on the estimator. The reduction of the currency portfolio risk through the shrinkage of the covariance and concentration matrixes, in theory, should lead to a decrease in the currency weights and therefore to a decrease of $(w_C\sigma_C)^2$ and σ_{PTF} . However, if the covariance between equities and indexes is negative, the decrease of w_C will lead to a lower diversification benefit, which can be crucial and even drive the increase of the total portfolio risk. Moreover, such a modification of the currency portfolio could change its performance and therefore the covariance between currencies and indexes, which can lead to both positive and negative results. As we can see, the dependence is not linear, hence just the risk reduction of the currency portfolio itself cannot guarantee the reduction of the combined portfolio.

This hypothesis could be also supported by the hedging portfolio weights presented in Table 5.7. The change is significant and sharp for some estimators, particularly for the SQP.

Currency	Optimal Hedge	Unitary Hedge	Universal Hedge	SQP	Glasso	Ledoit-Wolf
CHF	0,39	-0,08	-0,07	0,34	0,22	0,38
JPY	-0,26	-0,08	-0,07	-0,48	-0,22	-0,25
CAD	-0,52	-0,08	-0,07	-1,25	-0,47	-0,50
EUR	-0,85	-0,42	-0,34	-1,43	-0,33	-0,77
GBP	0,52	-0,08	-0,07	0,43	0,26	0,51
SEK	-0,46	-0,08	-0,07	-0,65	-0,35	-0,46
USD	0,51	-0,08	-0,07	0,08	0,04	0,44
Sum	**-0,67**	**-0,92**	**-0,74**	**-2,96**	**-0,85**	**-0,65**

Table 5.7. Hedging weights as of 30.04.22

The SQP maximization increases drastically the absolute values of the hedging weights. This is a surprising result, given the fact that shrinkage is expected to achieve the opposite - to shrink the trade sizes. The glasso estimator and the Ledoit-Wolf estimator perform as expected and achieve the desired position shrinkage. The hedging weights for the glasso estimator are reduced significantly while for LW the reduction is low to none for some of the currencies. However, the total sum of the hedges is lower only for the LW method. Furthermore, while the shrinkage of the glasso and LW seem to perform as expected on the individual hedging sizes, it still does not improve the risk profile of the portfolio, including tail risk. This again suggests that as less shrinkage is applied on the hedge, the better the hedging performance is.

While the shrinkage methods have proven to enhance the performance of the optimizers and reduce the risk when applied to the entire portfolio estimator, their application to the hedging element of the portfolio seems to be superfluous and even detrimental to the portfolio risk characteristics. We can therefore conclude that the application of shrinkage techniques does not provide any meaningful economical or statistical benefits to the investors. Some academics also support that hypothesis and argue against improvement of the MVOH. For instance, Lence (1995) states that his results suggest that the hedging research's recent emphasis on "better" MVHs has been a waste of resources. In another more recent paper, Dark (2005) argues that the literature should focus more on the assumptions underlying the conventional minimum variance hedge ratio (MVHR), rather than improving the techniques used to estimate it. While in those examples the researchers use different approaches for the improvement of the hedging ratios, they still support the hypothesis that the use of sophisticated techniques on the MVH proves to be redundant and unnecessary.

6. Conclusion

Appealed by the superior out-of-sample performance of different shrinkage methods for covariance and concentration matrix, both in portfolio construction and risk management, we applied those in the setting of the portfolio currency risk hedging.

The out-of-sample performance of the proposed estimators was not as expected when applied to the hedging element of our internationally diversified portfolio. We show that while the glasso and the Ledoit-Wolf estimators performed as expected on the individual elements of the hedge, shrinking the matrices and the corresponding hedge sizes, this did not result in total portfolio risk reduction or improvement of the out-of-sample portfolio risk profiles, including tail risk.

We hypothesize that this is mainly due to the structure of the minimum variance hedge which involves the currency concentration matrix and the currencies and equity covariance matrix. While shrinkage may be decreasing the estimation error, it also seems to disturb the balance and optimality of the minimum variance hedge. Furthermore, the structure itself, involving non-square matrices, complicates further the proper application of the presented shrinkage methods.

In conclusion, we do not find any substantial economical or statistical benefits in the application of shrinkage methods on the minimum variance hedge and believe the application should be avoided.

Appendix

Table 1: Data description

Asset	Currency	N	Time period
MSCI **Italy** Price Index	EUR	Train: 300 Out-of-sample test: 135	02.1986 - 04.2022 Train: 02.1987 - 01.2011 Out-of-sample test: 02.2011 - 04.2022
MSCI **France** Price Index	EUR		
MSCI **Germany** Price Index	EUR		
MSCI **Japan** Price Index	JPY		
MSCI **Netherlands** Price Index	EUR		
MSCI **Norway** Price Index	NOK - quote currency		
MSCI **Belgium** Price Index	EUR		
MSCI **Sweden** Price Index	SEK		
MSCI **Switzerland** Price Index	CHF		
MSCI **UK** Price Index	GBP		
MSCI **USA** Price Index	USD		
MSCI **Canada** Price Index	CAD		
EUR	EUR/NOK		
JPY	JPY/NOK		
SEK	SEK/NOK		
CHF	CHF/NOK		
GBP	GBP/NOK		
CAD	CAD/NOK		
USD	USD/NOK		

References

Banerjee, O., El Ghaoui, L., & d'Aspremont, A. (2008). Model Selection Through Sparse Maximum Likelihood Estimation for Multivariate Gaussian or Binary Data. *Journal of Machine Learning Research*, *9*(1), 485-516. https://www.jmlr.org/papers/volume9/banerjee08a/banerjee08a.pdf

BIEN, J., & TIBSHIRANI, R. J. (2011, December). Sparse estimation of a covariance matrix. *Biometrika*, *98*(4), 807-820. https://www-jstor-org.ezproxy.library.bi.no/stable/23076173?sid=primo&seq=1

Black, F. (1989, Jul. - Aug.). Universal Hedging: Optimizing Currency Risk and Reward in International Equity Portfolios. *Financial Analysts Journal*, *45*(4), 16-22. https://www-jstor-org.ezproxy.library.bi.no/stable/4479236?seq=1

CAMPBELL, J. Y., SERFATY-DE MEDEIROS, K., & VICEIRA, L. M. (2010, February). Global Currency Hedging. *The Journal of Finance*, *65*(1), 87-121. https://www-jstor-org.ezproxy.library.bi.no/stable/25656286?searchText=Global+Currency+Hedging&searchUri=%2Faction%2FdoBasicSearch%3FQuery%3DGlobal%2BCurrency%2BHedging%26so%3Drel&ab_segments=0%2Fbasic_search_gsv2%2Fcontrol&refreqid=fastly-default%3A498d82

Dark, J. (2005, July). A Critique of Minimum Variance Hedging. *Accounting research journal*, *18*(1), 40-49. https://www-emerald-com.ezproxy.library.bi.no/insight/content/doi/10.1108/10309610580000674/full/pdf?title=a-critique-of-minimum-variance-hedging

DeMiguel, V., Garlappi, L., & Uppal, R. (2009, May). Optimal versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy? *The Review of Financial Studies*, *2*(5), 1915-1953. https://www-jstor-org.ezproxy.library.bi.no/stable/30226017?searchText=DeMiguel+et+al.+2007&searchUri=%2Faction%2FdoBasicSearch%3FQuery%3DDeMiguel%2Bet%2Bal.%2B2007%26so%3Drel&ab_segments=0%2Fbasic_search_gsv2%2Fcontrol&refreqid=fastly-default%3A437d88ae6f

Dempster, A. P. (1972, March). Covariance Selection. *Biometrics*, *28*(1, Special Multivariate Issue), 157-175. https://www-jstor-org.ezproxy.library.bi.no/stable/2528966?searchText=Covariance+Selection%2C+Dempster+1972&searchUri=%2Faction%2FdoBasicSearch%3FQuery%3DCovariance%2BSelection%252C%2BDempster%2B1972%26so%3Drel&ab_segments=0%2Fbasic_search_gsv2%2Fcontrol&r

De Roon, F., Eiling, E., Gerard, B., & Hillion, P. (2012, October 17). *Currency Risk Hedging: No Free Lunch*. Social Science Research Network. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1343644

De Santis, G., & Gérard, B. (1998). How big is the premium for currency risk? *Journal of Financial Economics*, *49*(3), 375 - 412. https://www-sciencedirect-com.ezproxy.library.bi.no/science/article/pii/S0304405X98000294

Dumas, B., & Solnik, B. (1995, June). The World Price of Foreign Exchange Risk. *The Journal of Finance*, *50*(2), 445-479. https://www-jstor-org.ezproxy.library.bi.no/stable/2329415?seq=1

Eiling, E., & Gerard, B. (2015). Emerging Equity Market Comovements: Trends and Macroeconomic Fundamentals. *Review of Finance*, *19*(1), 1543 – 1585. https://bibsys-almaprimo.hosted.exlibrisgroup.com/primo-explore/fulldisplay?docid=TN_cdi_proquest_journals_1694727932&context=PC&vid=BI&lang=no_NO&search_scope=default_scope&adaptor=primo_central_multiple_fe&tab=default_tab&query=any,contains,Bruno%20%26%2

Friedman, J., Hastie, T., & Tibshirani, R. (2008, July). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432 - 441. https://academic.oup.com/biostatistics/article/9/3/432/224260?login=true

Glen, J., & Jorion, P. (1993, December). Currency Hedging for International Portfolios. *The Journal of Finance*, *48*(5), 1865-1886. https://www-jstor-org.ezproxy.library.bi.no/stable/2329070?seq=1

Goto, S., & Xu, Y. (2015, December). Improving Mean Variance Optimization through Sparse Hedging Restrictions. *The Journal of Financial and Quantitative Analysis*, *50*(6), 1415-1441. https://www-jstor-org.ezproxy.library.bi.no/stable/43862301?searchText=Improving+Mean+Variance+Optimization+through+Sparse+Hedging+Restrictions&searchUri=%2Faction%2FdoBasicSearch%3FQuery%3DImproving%2BMean%2BVariance%2BOptimization%2Bthrough%2BSparse%2BHe

Khoury, S. J., & Jorion, P. (1996). *Financial Risk Management: Domestic and International Dimensions*. Wiley. https://www.wiley.com/en-us/Financial+Risk+Management%3A+Domestic+and+International+Dimensions-p-9781557865915

Ledoit, O., & Wolf, M. (2004, February). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, *88*(2), 365-411. https://www-sciencedirect-com.ezproxy.library.bi.no/science/article/pii/S0047259X03000964

Ledoit, O. R. H. (1995). *Essays on risk and return in the stock market* [Thesis (Ph. D.)]. MIT Libraries. https://dspace.mit.edu/handle/1721.1/11875

Lence, S. H. (1995, May). The Economic Value of Minimum-Variance Hedges. *American Journal of Agricultural Economics*, *77*(2), 353-364. https://www-jstor-org.ezproxy.library.bi.no/stable/1243545?sid=primo&origin=crossref&seq=1

Michaud, R. O. (1989, January - February). The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal*, *45*(1), 31-42. https://www-jstor-org.ezproxy.library.bi.no/stable/4479185?seq=1

Millington, T., & Niranjan, M. (2017, October 26). *Robust Portfolio Risk Minimization Using the Graphical Lasso*. Springer. https://link.springer.com/chapter/10.1007/978-3-319-70096-0_88

Perold, A. F., & Schulman, E. C. (1988, May - June). The Free Lunch in Currency Hedging: Implications for Investment Policy and Performance Standards. *Financial Analysts Journal*, *44*(3), 45-50. https://www-jstor-org.ezproxy.library.bi.no/stable/4479112?searchText=Free+lunch+currency&searchUri=%2Faction%2FdoBasicSearch%3FQuery%3DFree%2Blunch%2Bcurrency%26so%3Drel&ab_segments=0%2FSYC-6427%2Fcontrol&refreqid=fastly-default%3A58e2a75dad5219c348474bf7

Roll, R. (1992). Industrial Structure and the Comparative Behavior of International Stock Market Indices. *The Journal of Finance*, *Vol. 47*(No. 1 (Mar., 1992)), 3-41. https://www-jstor-org.ezproxy.library.bi.no/stable/2329089?seq=1

Solnik, B. H. (1974). Why Not Diversify Internationally Rather than Domestically? *Financial Analysts Journal*, *30*(No. 4 (Jul. - Aug., 1974)), 48-52+54 (6 pages). https://www.jstor.org/stable/4529718?seq=1

Yuan, M., & Lin, Y. (2007, March). Model Selection and Estimation in the Gaussian Graphical Model. *Biometrika*, *94*(1), 19 - 35. https://www-jstor-org.ezproxy.library.bi.no/stable/20441351?sid=primo&seq=1