ELIT
Economic Laboratory Transition
Research Podgorica

Montenegrin Journal of Economics

For citation:

Nhat, N.M. (2025), "Covariance Matrix Estimation for Portfolio Selection: Linear and Non-Linear Shrinkage Methods on the Frontier Market", *Montenegrin Journal of Economics*, Vol. 21, No. 4, pp. 217-227.

Covariance Matrix Estimation for Portfolio Selection: Linear and Non-Linear Shrinkage Methods on the Frontier Market

NGUYEN MINH NHAT[1]

[1], Lecturer, Faculty of Banking Ho Chi Minh University of Banking (HUB), Vietnam, E-mail: nhatnm@hub.edu.vn

ARTICLE INFO

Received July 30, 2024
Revised from August 29, 2024
Accepted September 30, 2024
Available online October 15, 2025

JEL classification: C14, C58, G14, G41

DOI: 10.14254/1800-5845/2025.21-4.17

Keywords:

Non-linear Shrinkage estimators,
custom-tailored Shrinkage,
covariance matrix estimation,
portfolio selection

ABSTRACT

The development of a stable, secure, and efficient stock market in Vietnam plays a crucial role in enhancing risk resilience, ensuring a well-structured market, and serving as a platform to attract substantial medium and long-term capital for the economy. Therefore, the study aims to analyze the effectiveness of linear and non-linear shrinkage estimators for portfolio selection in the Vietnamese stock market. Comparing various estimators, including custom-tailored linear Shrinkage, non-linear Shrinkage, and generic linear Shrinkage. The results reveal that the non-linear shrinkage estimator does not always give superior results compared to the generic linear shrinkage estimator, especially when dealing with low numbers of stocks. The custom-tailored linear shrinkage estimator remains a challenging benchmark that the non-linear shrinkage estimator needs to overcome. These findings contribute to further empirical evidence on the effectiveness of shrinkage estimators for portfolio selection in a frontier market. Finally, the author proposes several policy implications for the stock market to promote and create necessary guidance for Vietnamese businesses and economic organizations to apply international financial standards, helping to improve modern management and corporate governance capacity in line with the world's development and progress trends.

INTRODUCTION

The covariance estimator plays an essential role in portfolio optimization and can determine the success or failure of a portfolio investment strategy. Initially, the sample covariance estimator can support the investors in estimating the covariance matrix, but its effectiveness in practice is not high, and there are problems in the case of high–dimensional data (John & Mieldzioc, 2020). The researchers try to propose many different methods of estimating the covariance matrix to solve this problem, notably the Shrinkage of covariance estimator first introduced (Ledoit & Wolf, 2003a, 2003b, 2004). The estimator has shown its effectiveness when applied in many different fields, especially in selecting the optimal asset portfolio (DeMiguel et al., 2013).

The shrinkage methods, as in their name, describe a collection of computing approaches that aim to eliminate the most minor impact factors from the whole ones. It takes into account a sample covariance matrix (SC) and a structural target in the attempt to shrink the noise and bias of the estimated covariance matrix. The original shrinkage method proposed by O. Ledoit & M. Wolf (2003b) includes a balance factor to adjust the level of tradeoff of the estimation called shrinkage intensity. Their works have been

proven to be consistent with different sizes of high-dimensional portfolios. O. Ledoit & M. Wolf (2022) published a journal paper summarizing the applications of the shrinkage estimator over the past 15 years and describing in detail the theoretical framework and instructions on how to use this method for covariance matrix estimation. In their paper, a list of aspects was discussed carefully such as asset optimization together with some new shrinkage-based approaches for the means of sample returns and covariance matrices, and they also consider how to derive the weights in the asset. In order to improve the performance of small size portfolios with small estimation sample size, B. Candelon et al. (2012) also introduced shrinkage-based estimators that can be beneficial in realistic asset management. Even more practical, A. Kourtis et al. (2012) suggested to use two non-parametric methods to modify the shrinkage technique that can be applied directly to the inverse covariance matrix. Their method especially can be useful to resolve the case of high transaction costs, short-sale conditions, and singular covariance matrices, while keeping lower risks and high risk-adjusted returns.

O. Ledoit & M. Wolf (2022) had shown that there are two important approaches to generate a reliable shrinkage covariance matrix for portfolio selection. The first approach requires to select manually the target for Shrinkage based on the experiences of experts to come closely to the real covariance matrix but ensuring the fewest possible number of parameters. Clearly, this approach requires some sorts of knowledge on the feature space of the true covariance matrix. This knowledge can be represented as the asset returns, a positive average covariance or a positive average correlation. Thus, this approach is called custom-tailored target-based linear Shrinkage. However, this can be fulfilled in many cases of portfolios, and that led Ledoit and Wolf considered a second approach based on a non-linear structure of shrinkage estimator for covariance matrix estimation. This approach imposes the relationship of empirical covariance matrix and the generic shrinkage target. The estimation process is hence modeled purely by mathematical formulation and does not require any specific experiences of the experts as in their previous study. Also in this work, they performed a deep analysis to prove the improvement of non-linear shrinkage approach over the linear one.

To our best knowledge, there are plenty of research so far regarding to the comparison, both pros and cons of two approaches: linear and non-linear shrinkage methods in the world. However, there is almost no clear demonstration of this analysis for Vietnam's stock market. In this paper, we conduct empirical analysis on the effectiveness of versions of shrinkage estimators and answer the question whether the modern non-linear shrinkage estimator can beat the traditional linear estimators, especially the linear shrinkage ones built on some sorts of knowledge of the finance experts. Moreover, our experimental results are expected to show potential achievement in portfolio selection in Vietnam's stock market and contribute to further empirical evidence on the effectiveness of the shrinkage estimators for optimal portfolio selection.

The remaining contents of the research paper will be presented in the following order: Section 2 is related studies, section 3 is theoretical framework, section 4 is data and methodology, section 5 is empirical result, and the conclusion in section 6.

1. LITERATURE EMPIRICAL REVIEW

1.1 Related to studies

We all know that common research in modern portfolio management mostly approaches from quantitative perspectives. Specifically, people focus on the mathematical modeling of a portfolio based on the objectives of optimization problems. Although it sounds simple, plenty of significant researches have been raised from this idea and they can prove their impact to modern portfolio studies. Among them, it can be figured out the framework proposed by O. Ledoit & M. Wolf (2003b) which resolves the optimization problem by shrinkage methodology. In their first study, Ledoit and Wolf believed that using the first characteristic for pricing assets in the CAPM as a shrinkage target leads to an efficient approximation of the covariance matrix of a portfolio since, from their practical experiences, stock returns have at least a factor represented for model structure.

In an extended version, O. Ledoit & M. Wolf (2004) dug into the second feature as called the constant correlation model (CC) and used it as the replacement of shrinkage target taking into account the prior belief of positive average correlation of all stocks when considered at the same time. Both of their results have shown that shrinkage target towards single index model (SSI) and shrinkage target towards constant correlation model (SCC) are commonly outperform the traditional sample covariance matrix estimation methods. However, in practice, SCC often provides best selection of portfolio for small sample size (generally less than 100 stocks). The SSI, on the other hand, gives best selection results with large sample size (larger than 225 stocks). Rely on Ledoit and Wolf framework, Schafer & Strimmer (2005) considering covariance matrices from sparse genomic data, proposed the third shrinkage target could get the best estimation for both small and large scale matrices.

Next, in 2012, based on the observation that the average covariance of hedge fund returns has a positive value, B. Scherer & K. Winston (2011) estimated the covariance matrix by applying the shrinkage target which shows an important application of shrinkage methods. It can be observed in many practical uses case, using multiple shrinkage targets may push the overall performance of estimated covariance matrix instead of using only linear Shrinkage designed with a single target. L. Garlappi et al. (2007) revealed a multiple prior targets model without ambiguity to examine the normality of parameter and model uncertainty for portfolio management. They showed that their proposed approach for investment selection is stable across different time periods and achieve higher out-of-sample Sharpe ratio. With the same approach of using multiple priors, F. Lutgens & P. Schotman (2010) studied the robust optimization based on different experiences or knowledge from experts of the field. The content is that they considered the distinct priors from a finite set of hypotheses, with the prior probabilities are attached to each of the priors, counted on the assumption that the aversion to ambiguity of a robust investor. If there are significant estimation errors, it can be proven the multi-prior models are useful for portfolio selection in practice.

In the case of rebalance the uncertainty through a weighted averaging, J. Fan et al. (2016) considered the targeting multiple structure covariance matrices by which we can discover more information to facilitate the shrinkage intensities estimation. From the experimental results, people can see the multi-target Shrinkage is superior to the single one in out-of-sample performance, in the sense that it can reduce significantly constraints of conditional expected returns and the portfolio variance for different portfolio sizes, types. Moreover, it gets higher information ratios purely for medium to small sized portfolio structures.

Recently, there have been several approaches inspired by data-driven perspectives. Among them, it can be figured out J. Tong et al. (2018), in which the authors considered the leave-one-out cross-validation (CV) approach to select the data-driven shrinkage factor. The applied loss is the Frobenius norm of the estimation error, and their learning method is aiming to minimize this loss. Tong's method can provide the estimators with low-complexity and can be developed on multi-targets Shrinkage. The numerical expressions demonstrate the effectiveness of the proposed approach for obtaining the oracle approximating performance of the estimated covariance matrix, and it can enhance the effectiveness of some other array signal processing problems. T. Tran et al. (2020) suggested a combination of the original shrinkage-based estimation framework and evaluation technique of CV to derive the optimal shrinkage coefficient with different uncertainty. Their illustration of real data in the Vietnamese stock market can prove the potentials of data-driven based approach with more accurate evaluation in a real-time manner.

There is another branch of research which does not rely on the known features or characteristics of the true covariance matrix, but we can still select a suitable covariance matrix as accurate as the linear shrinkage approach to a custom-tailored target. This research direction was pioneered by O. Ledoit & M. Wolf (2004), which appeared in a thorough study of how to combine the empirical covariance matrix and an identity matrix in the estimation of covariance matrix. From the empirical results, the generic linear Shrinkage to identity matrix was shown its superior than other traditional estimators, and the larger size of portfolio is, the more pronounced was the difference. In O. Ledoit and M. Wolf (2012, 2015), the authors switched to the extension of their work via focusing on the non–linear structures of the shrinkage estimators for covariance matrix estimation problem. Their approach can be seen as an indirect NL shrinkage that aim to recover the population eigenvalues with the help of Quantized Eigenvalues Sam-

pling Transform (QuEST) function. Then a plenty of extensive research on using Monte Carlo simulation revealed that indirect NL shrinkage is generally better than the generic linear Shrinkage as we mentioned before, and often useful in the case of a large margin.

A study of C. Lam et al. (2017) introduced to another non–linear shrinkage method, which is called NERCOME, which devides the whole sample into two critical parts: one is for estimating the eigenvectors, and the other one aims to estimate those eigenvalues associated with these eigenvectors. It can be observed that the main attraction of NERCOME is that it is efficient, decisively simpler and more transparent, hence supplying an basement for the elegant mathematical formulation of the indirect method of the QuEST function. O. Ledoit & M. Wolf (2020), on the other hand, struggled with a fresh approach, regarding to the analytical (or direct) non–linear Shrinkage, which took advantages of a nonparametric kernel density estimation (DKE) of the limiting spectral signals of sample eigenvalues and its Hilbert transformation. Their proposal has been proven effective both in theoretical and practical perspectives, which includes the best qualities of three approaches already mentioned above: the accuracy of the considered QuEST function, the convergence speed of linear Shrinkage, and the intrinsic transparency of NERCOME. Besides, there are also some studies aimed at improving the efficiency of the NL estimator, such as the study of C. Bongiorno & D. Challet (2023) has demonstrated that using the RIE to calculate the covariance matrix eigenvalues gives much better results than using the Oracle.

1.2 Portfolio Optimization

There are many ways to achieve such an optimum portfolio and one of a quantitative approach is to search for a portfolio with a minimum variance. In the case we have N tickers of stocks in our investment with the stock weights $m=(m_1, m_2, \dots, m_N)$, the summation of weights should be equal to $1\left(\sum_{i=1}^{N} m_i=1\right)$, where $m_i>0$ means that short selling will not exist. The portfolio selection problem should be modeled as follow:

$$\begin{aligned} &argmin \quad m^{\intercal}\Sigma m \\ &\text{s.t.} \quad m^{\intercal}1=1 \end{aligned}$$

¿ ¿

Here, we denote 1 and Σ respectively a vector of all ones and covariance matrix of N stocks as we mentioned above. Due to the convexity of the objective function (1), the problem can be solved and derived

¿

We should note that the theoretical solution (2) involves the precision matrix, which is normally replaced by the sample covariance matrix (SC), hence it may lead to the scenario of singularity. It means that sample covariance is generally ill-posed, which is invertible in most of practical use cases, particularly in the circumstance of high-dimensional portfolios. Therefore, using shrinkage estimators' approach is expected to provide feasibly numerical solution for the equation (2), and hence, we can output the stocks weights for our investment.

1.3 Linear Shrinkage to generic target (GL)

Among the linear shrinkage approaches to generic target, the Shrinkage to identity matrix is quite popular, which consists of a weighted combination of the SC and the identity matrix (IM). In this approach, the identity matrix, with all elements in diagonal are equal to one and zero elsewhere, is known as a target matrix. Although the shrinkage target matrix does not possess any known feature of the real covariance matrix and high biased, it still contains one of important properties, which is no variance. So, the Shrinkage to identity matrix can generate a covariance matrix with more stable than the one of sample covariance matrix.

Assuming that Σ is a true covariance matrix, S denotes a sample covariance matrix, I is an identity matrix, and $\delta^{¿}$ is shrinkage intensity $(0 \le \delta^{¿} \le 1)$. The generic shrinkage covariance matrix $\Sigma^{¿}$ is estimated as follows:

$$\Sigma^{¿} = \delta^{¿} \mu I + (1 - \delta^{¿}) S$$

with $\delta^{¿} = \frac{\beta^2}{\alpha^2}$.

Wherein: $\mu = \langle \Sigma, I \rangle; \beta^2 = E[\| S - \Sigma \|^2]; \sigma^2 = E[\| S - \mu I \|^2]$.

1.4 Linear Shrinkage to a custom-tailored target (CTL)

A custom–tailored target is a shrinkage target established from the application of specific knowledge. O. Ledoit & M. Wolf (2020) said that: "The financial covariance matrices typically have some stylized features that can be exploited in devising an improved shrinkage target compared to using (a multiple of) the identity matrix". This is known as an expert-based idea for a generic target. In practice, to achieve an efficient custom–tailored target, we are not only interested in how well it fits the true covariance matrix but also whether the number of parameters it generates is small or not.

In this paper, we use the covariance matrix of single index model (SI) and the one of constant correlation model (CC) as the custom–tailored targets of shrinkage estimation, respectively named Shrinkage to single-index model (SSI) and Shrinkage to constant correlation model (SCC). The corresponding shrinkage target matrix from SI generally has a factor represented for model structure that reflects that market return has a significant impact on the stocks' return of selected portfolios. Meanwhile, the CC target is estimated assuming that the correlation between all pairs of stocks is no different and is equal to the average correlation level (Ledoit & Wolf, 2020).

The CTL is defined as a linear combination between a SC and a custom–tailored target. The optimal shrinkage coefficient $\delta^{¿}$ can be used to score the trade-off between these two covariance matrices. The value of shrinkage coefficient will be from zero to one, so the closer the shrinkage coefficient approaches to a value of 1, the greater the impact of custom–tailored target on the estimated covariance matrix.

Based on the primary study of O. Ledoit & M. Wolf (2003b, 2004), the covariance matrix from the CTL is calculated as follows:

$$\hat{\Sigma} = (1 - \delta^{¿}) S + \delta^{¿} T .$$

With:

$$\delta^{¿} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \left[\mathrm{Var}(s_{ij}) - \mathrm{Cov}(t_{ij}, s_{ij}) \right]}{\sum_{i=1}^{N} \sum_{j=1}^{N} \left[Var(t_{ij} - s_{ij}) + (\theta_{ij} - \sigma_{ij})^2 \right]}$$

herein: σ_{ij}, s_{ij} and t_{ij} are elements of Σ, S and T respectively; and $\theta_{ij} = E(f_{ij})$ and $\sigma_{ij} = E(s_{ij})$.

1.5 Non-linear shrinkage (NL)

In this paper, we apply the analytical NL estimator introduced by O. Ledoit & M. Wolf (2020) for estimating the sample covariance matrix. Recent studies have shown the superiority of this method compared to other non-linear shrinkage estimators when it is able to effectively balance three issues in estimating the covariance matrix: level of accuracy, calculation speed and transparency. Estimating the covariance matrix using the analytical non–linear shrinkage estimator can be described through the following formula:

$$\hat{\Sigma}^{\bullet} := V \hat{\Delta}^{\bullet} V'$$

with

$$\hat{\Delta}^{\bullet} := \mathrm{Diag}\left(\hat{\Phi}^{\bullet}(\lambda_1), \ldots, \hat{\Phi}^{\bullet}(\lambda_N) \right)$$

where

¿

The fundamental distinction of NL and GL shrinkage is the way eigenvalues are handled. In the GL estimator, the eigenvalues of sample are passed towards their grand mean, whereas according to the NL estimator, some eigenvalues of the sample may be removed out of their grand mean and passed towards local centroids of attraction. However, the NL method is useful to reduce the overall fluctuation if we compare with the other estimator. Hence, theoretically, it has better adaptation than the generic linear Shrinkage in estimating the covariance matrix.

2. METHODOLOGY AND DATA

We conduct experiments on shrinkage estimators on the Vietnamese stock market. This is considered one of the emerging financial markets in Southeast Asia with rapid development recently. Input data is daily stock prices of more than 400 stock codes listed on the Ho Chi Minh City Stock Exchange, one of the two largest trading floors in Vietnam. To ensure the quality of research results, the stock codes of the companies being researched must ensure the following requirements: (i) Have been listed on the exchange for at least 2 years, (ii) Ensure stock liquidity with an average trading volume of 10,000 shares per week, and (iii) Eliminate penny stocks priced under $0.20 per share. The research period is about 10 years from 2011 to 2021, this is considered a period of stable development of the Vietnamese stock market.

To examine the effectiveness of the different shrinkage estimators, a backtesting procedure is applied based on the research results of Tran et al. (2020) and described specifically in the study of N. Nguyen et al. (2020). The rolling horizon procedure, with a time window of 104 weeks that moves forward a week each iteration, is used in the backtesting system to calculate the performance metrics of portfolios. The portfolios will be rebalanced weekly, this is also consistent with the features of the Vietnam's stock market when the trade settlement cycle is date T+2. The in-sample period will start from January 1, 2011, to December 31, 2012, corresponding to 104 weeks used to initialize the investment portfolio. The out-of-sample period will start from January 1, 2013, to January 1, 2021, corresponding to 416 weeks, used to test the effectiveness of the estimation methods.

Moreover, the metrics used to evaluate the performance of shrinkage estimators are quite comprehensive; They not only include indicators of profit (Annual Return - AR) and risk (Annual Volatility - AV) of the portfolio but also criteria for considering other aspects of the portfolio such as: (i) Correlation ratio between profit and risk (Sharpe ratio – SR), (ii) The level of stability of portfolio (Portfolio Turnover – PT), (iii) Maximum loss of portfolio (ML), (iv) The probability of occurring the worst case of possible loss over a target threshold with a given confidence level (Value at risk – VaR), and (v) Abnormal return of portfolio (Jensen's Alpha – JA).

Additionally, to better reflect the practicality of the trading strategy, transaction costs are also taken into account when calculating the performance metrics. We apply a transaction fee of 0.25% of the total transaction value for each purchase or sale of shares in the backtesting system. This is the fee being applied on the Vietnam's stock market. Besides, to be fair, we not only apply the estimators on a single portfolio (N = 400 shares as mentioned before), but we also test these estimators' methods on the different groups of portfolios with the predefined stock numbers 100, 200, 300, 400, respectively. The distribution of stocks weights into portfolios depends on the market capitalization of those stocks in the considered period. For instance, taking N = 100 means that the investors will hold 100 securities in their portfolio, with the possibly largest market capitalization, and similar to N = 200, 300, 400.

3. EMPIRICAL RESULTS

The Vietnamese stock market experienced a volatile 2023; however, it still ended on a positive note. 2024 is expected to continue to be a promising year with global economic recovery and strong support

policies from the Government. In 2023, it is a year of overcoming difficulties for the stock market and promising growth prospects in 2024. Amidst a year marked by unpredictable changes in the global economy and geopolitical factors, Vietnam has demonstrated resilience by maintaining a commendable GDP growth rate of 5.05%. Specifically, the cumulative market value of the Vietnamese stock market exceeded 240 billion USD by the conclusion of 2023, which accounted for 56.4% of the country's GDP. Notably, the Ho Chi Minh City Stock Exchange (HOSE) alone contributed 186 billion USD to this total. Foreign investors divested a total of 985.8 million shares, resulting in a net outflow of VND 24,830.9 billion (about USD 1 billion). Previously, in 2022, foreign investors net bought VND 26,674 billion on the HOSE floor. Thus, after strong net buying in 2022, foreign investors increased net selling in 2023. In the first three months of 2024, the stock market also had a positive start when, at the end of the trading session on March 25, 2024, the VN-Index reached 1,267.9 points, with the matched transaction value on the HOSE floor reaching VND 27.2 trillion.

The main factors driving the home economy include initiatives to accelerate the distribution of public investment, measures to boost domestic consumption, tax reductions, and pay hikes in both the public and private spheres. Furthermore, the endorsement by the National Assembly of many amended laws, including the Land Law, the Law on Credit Institutions, the Housing Law, etc., would provide a secure legal framework, providing the groundwork for development in the upcoming era. The decisive measures implemented by management agencies to address challenges in the real estate market and corporate bond market have yielded favorable outcomes, while maintaining investor confidence. According to the World Bank's estimates, the Vietnamese stock market could attract an additional US$25 billion from foreign investors by 2030 after the market is upgraded. For the economy, the development of the financial market contributes to reducing the economy's dependence on bank credit, thereby minimizing potential credit risks. In addition, when the stock market is successfully upgraded, the role and business opportunities for commercial banks and foreign bank branches are also expanded in both the primary and secondary markets. Commercial banks and foreign bank branches will benefit from the provision of financial services related to issuance, direct investment, distribution and issuance of guarantees as well as products such as bond trading, derivatives, lending, deposits, securities custody, foreign exchange, etc. Therefore, the development of the medium and long-term stock market, contributing to the goal of attracting foreign capital flows in a more sustainable manner.

Table 1. The effectiveness of the estimators on the top 100 largest stocks by market capitalization during the out-of-sample testing period

Approaches	*Estimators*	*AR*	*AV*	*SR*	*PT*	*ML*	*VaR*	*JA*
Generic	GL	12.52	10.36	0.80	3.60	-32.28	-1.26	4.06
Custom-tailored	SSI	12.58	10.45	0.79	3.99	-29.57	-1.27	4.19
	SCC	14.20	10.44	0.93	3.44	-30.59	-1.26	5.57
Non-linear	DKE	11.60	10.29	0.73	3.98	-32.64	-1.25	3.25

Source: Author's calculations

Table 1 showed that the SCC estimator leads with the highest Jensen's Alpha is 5.57, indicating superior performance over the expected return. The SSI estimator is 4.19 next, followed by GL is 4.06 and DKE is 3.25. Besides, the SCC estimator in the custom-tailored approach appears to be the most effective overall, with the highest average return, Sharpe Ratio, and Jensen's Alpha, combined with relatively low volatility and portfolio turnover. The GL and SSI estimators also perform well but are slightly behind SCC in several metrics. The DKE estimator, while having the lowest volatility, lags in average return, Sharpe Ratio, and Jensen's Alpha, making it less effective in comparison.

Table 2. The effectiveness of the estimators on the top 200 largest stocks by market capitalization during the out-of-sample testing period

Approaches	*Estimators*	*AR*	*AV*	*SR*	*PT*	*ML*	*VaR*	*JA*
Generic	GL	16.18	8.79	1.31	4.15	-20.8	-1.05	8.25
Custom-tailored	SSI	17.20	8.84	1.40	4.78	-16.86	-1.05	9.35
	SCC	19.01	8.95	1.55	3.91	-14.98	-1.06	10.86

Non-linear	DKE	16.60	8.77	1.37	3.68	-19.66	-1.04	8.74

Source: Author's calculations

Table 2 showed that the SCC estimator leads with the highest Jensen's Alpha is 10.86, indicating superior performance over the expected return. The SSI estimator is 9.35 next, followed by DKE is 8.74 and GL is 8.25. Besides, the SCC estimator in the custom-tailored approach appears to be the most effective overall, with the highest average return, Sharpe Ratio, and Jensen's Alpha, combined with relatively low volatility and portfolio turnover. The SSI and DKE estimators also perform well but are slightly behind SCC in several metrics. The GL estimator, while performing adequately, lags behind the others in key metrics such as average return, Sharpe Ratio, and Jensen's Alpha, making it less effective in comparison.

Table 3. The effectiveness of the estimators on the top 300 largest stocks by market capitalization during the out-of-sample testing period

Approaches	*Estimators*	*AR*	*AV*	*SR*	*PT*	*ML*	*VaR*	*JA*
Generic	GL	18.36	8.51	1.57	4.35	-16.88	-1.00	10.37
Custom-tailored	SSI	19.01	8.50	1.63	4.92	-13.93	-1.00	11.12
	SCC	20.26	8.73	1.70	4.03	-13.68	-1.03	12.04
Non-linear	DKE	19.03	8.53	1.64	3.89	-17.92	-1.00	11.11

Source: Author's calculations

Table 3 showed that the SCC estimator leads with the highest Jensen's Alpha is 12.04, indicating superior performance over the expected return. The SSI estimator is 11.12 and DKE is 11.11 are also strong, with GL is 10.37 being lower. Besides, the SCC estimator in the custom-tailored approach appears to be the most effective overall, with the highest average return, Sharpe Ratio, and Jensen's Alpha, combined with relatively low volatility and portfolio turnover. The DKE and SSI estimators also perform well but are slightly behind SCC in several metrics. The GL estimator, while performing adequately, lags behind the others in key metrics such as average return, Sharpe Ratio, and Jensen's Alpha, making it less effective in comparison.

Table 4. The effectiveness of the estimators on the top 400 largest stocks by market capitalization during the out-of-sample testing period

Approaches	*Estimators*	*AR*	*AV*	*SR*	*PT*	*ML*	*VaR*	*JA*
Generic	GL	19.74	8.36	1.74	4.28	-17.95	-0.98	11.74
Custom-tailored	SSI	20.41	8.34	1.80	4.86	-13.83	-0.98	12.51
	SCC	21.25	8.65	1.81	3.99	-13.23	-1.01	12.99
Non-linear	DKE	20.38	8.38	1.80	3.87	-18.59	-0.98	12.47

Source: Author's calculations

Table 4 showed that the back-testing results of covariance matrix estimators on out–of–sample, we can draw some important points as follows: both custom–tailored linear shrinkage (CTL) and non-linear shrinkage (NL) give much superior results than the generic linear Shrinkage (GL) on most portfolio performance metrics. In which, the CTL portfolios outperform the GL portfolios with five out of seven evaluation criteria including average annual return, Shape ratio, portfolio turnover, MDD and Alpha coefficient; meanwhile the NL portfolios show complete superiority over the GL portfolios with six out of seven criteria, except for MDD criterion. This experimental result is generally consistent with the most recent study (Ledoit & Wolf, 2020).

However, there is an interesting point that the NL only produces a clear superiority to the GL when the number of stocks in the portfolio (N) tends to increase; conversely when N is small, the NL is not as efficient as the GL. Specifically in this study, when N is equal 100, the average annual return, Sharpe ratio, daily portfolio turnover, MDD and Jensen’s Alpha of generic linear Shrinkage have values of 12.52%, 0.8, 3.6%, 32.28%, and 4.06% that are better than the ones of non-linear Shrinkage with values

of 11.6%, 0.73, 3.98%, 32.64%, and 3.25% respectively. The GL is only inferior to the NL in two evaluation criteria that are average annual volatility and VaR in this case. On the other hand, when N is larger than 200, the non-linear Shrinkage outperforms the generic linear Shrinkage on most of considered performance metrics. This superiority is even more evident when the number of stocks in the tested portfolio increases rapidly from 200 to 400 stocks.

The custom-tailored linear shrinkage portfolios are likely to generate superior returns compared to the non-linear shrinkage portfolios, while the non-linear shrinkage portfolios tend to be less risky than the custom–tailored linear Shrinkage. However, if we consider return of portfolio compared with its risk, the average Shapre ratio of CTL portfolios give higher values than the one of NL portfolios. In particular, the average annual return and Jensen's Alpha on four tested portfolios of CTL are about 18% and 9.9% which are much better than the ones of NL with the values of 16.9% and 8.9% respectively. Meanwhile, the other performance metrics of CTL, which reflect the level of portfolio risk such as average annual volatility, daily turnover portfolio and VaR with the values of 9.1%, 4.24%, and 1.08%, are not as good as the NL's with corresponding values of 8.9%, 3.8%, and 1.06%. Although the CTL portfolios are riskier than the NL portfolios, if we compare the portfolio's return per unit of risk, we can conclude that the CTL portfolios still outperform the NL portfolios. This is clearly demonstrated when the average Sharpe ratio of CTL portfolios on four back-testing portfolios are much higher than that of NL portfolios, 1.5% versus 1.3%.

The difference in evaluation results between the custom-tailored linear shrinkage approach and the non-linear shrinkage approach has decreased sharply when the number of stocks in the portfolio (N) increases dramatically. For example, when N is equal to 100, the superiority of CTL compared to NL on the criteria annual return, SR and Alpha is 2.6%, 0.2, and 2.32% respectively, however when N approaches 400, the difference is This difference is only 0.87%, 0.01 and 0.52%.

CONCLUSION

Based on the analysis above, there is plenty of research so far regarding the comparison, both pros and cons of two approaches: linear and non-linear shrinkage methods in the world. However, there is almost no clear demonstration of this analysis for the financial markets of Southeast Asian countries. In this article, we conduct practical experiments on the applicability of these two research directions. The experimental research results have discovered two very interesting points. First, the non-linear shrinkage (NL) estimator does not always give superior results compared to the generic linear Shrinkage (GL) estimator; In case the number of stocks in the portfolio is low (N=100), the NL is not as effective as the GL on most portfolio performance metrics. Second, the custom–tailored linear shrinkage (CTL) estimator is still the tough benchmark that the NL estimator needs to overcome. The CTL has shown its strength when it prevails in all measures of the portfolio's ability to generate outstanding returns such as annual return, Sharpe ratio or Jensen's Alpha compared to the NL even though the NL tends to create portfolios with low volatility. These findings show that the academic researchers and practitioners still have a lot of work to do if the NL approach is to become more popular and widely applied in the financial markets. Moreover, our experimental results show potential achievement in portfolio selection in a frontier market such as the Vietnamese stock market and contribute to further empirical evidence on the effectiveness of the shrinkage estimators for optimal portfolio selection.

The limitations and further research. The study uses stock data from the Ho Chi Minh City Stock Exchange over a ten-year period. This may not fully capture the broader market dynamics or rare market events that could affect the generalizability of the findings. This limits the applicability of the results to other emerging or developed markets with different characteristics and investor behaviors. The study concludes that non-linear shrinkage estimators do not always outperform linear ones, especially with smaller portfolios. However, the research does not extensively explore why this is the case or provide a detailed analysis of the underlying reasons for this performance discrepancy. The use of a rolling window for backtesting may introduce bias, as it assumes that past performance is a reliable indicator of future results. This could lead to overfitting the model to historical data. Therefore, further research suggestions are expand the dataset to include a more diverse range of stocks, including those with lower liquidity or smaller market capitalizations, to better understand their impact on portfolio selection and covariance

matrix estimation. Conduct similar studies in different markets (emerging, developed) to compare the effectiveness of linear and non-linear shrinkage estimators across varying market conditions and investor behaviors.

ACKNOWLEDGEMENT

The author thanks all lecturers and managers of Ho Chi Minh University of Banking (HUB). This article is supported and funded by HUB, Vietnam.

REFERENCES

Bongiorno, C., Challet, D. (2023), "Non-linear shrinkage of the price return covariance matrix is far from optimal for portfolio optimization", *Finance Research Letters*, Vol. 52, No. 3, pp. 1-15. https://doi.org/ 10.1016/j.frl.2022.103383

Candelon, B., Hurlin, C., Tokpavi, S. (2012), "Sampling error and double shrinkage estimation of minimum variance portfolios", *Journal of Empirical Finance*, Vol. 19, No. 4, pp. 511-527. https://doi.org/10. 1016/j.jempfin.2012.04.010

DeMiguel, V., Martin-Utrera, A., Nogales, F. J. (2013), "Size matters: Optimal calibration of shrinkage estimators for portfolio selection", *Journal of Banking & Finance*, Vol. 37, No. 8, pp. 3018-3034. https://doi.org/10.1016/j.jbankfin.2013.04.033

Fan, J., Liao, Y., & Liu, H. (2016), "An overview of the estimation of large covariance and precision matrices", *The Econometrics Journal*, Vol. 19, No. 1, pp. 1-32. https://doi.org/10.1111/ectj.12061

Garlappi, L., Uppal, R., Wang, T. (2007), "Portfolio selection with parameter and model uncertainty: A multi-prior approach", *The Review of Financial Studies*, Vol. 20, No. 1, pp. 41-81. https://doi.org/10. 1093/rfs/hhl003

John, M., Mieldzioc, A. (2020), "The comparison of the estimators of banded Toeplitz covariance structure under the high-dimensional multivariate model", *Communications in Statistics-Simulation and Computation*, Vol. 49, No. 3, pp. 734-752. https://doi.org/10.1080/03610918.2019.1614622

Kourtis, A., Dotsis, G., Markellos, R. N. (2012), "Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix", *Journal of Banking & Finance*, Vol. 36, No. 9, pp. 2522-2531. https://doi. org/10.1016/j.jbankfin.2012.05.005

Lam, C., Feng, P., Hu, C. (2017), "Non-linear shrinkage estimation of large integrated covariance matrices", *Biometrika*, Vol. 104, No. 2, pp. 481-488. https://doi.org/10.1093/biomet/asx021

Ledoit, O., Wolf, M. (2003a), "Honey, I shrunk the sample covariance matrix", pp. 1-22. http://dx.doi.org/10.2139/ssrn.433840

Ledoit, O., Wolf, M. (2003b), "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection", *Journal of Empirical Finance*, Vol. 10, No. 5, pp. 603-621. https://doi.org/10.1016/S0927-5398(03)00007-0

Ledoit, O., Wolf, M. (2004), "A well-conditioned estimator for large-dimensional covariance matrices", *Journal of Multivariate Analysis*, Vol. 88, No. 2, pp. 365-411. https://doi.org/10.1016/S0047-259X(03)00096-4

Ledoit, O., Wolf, M. (2012), "Non-linear shrinkage estimation of large-dimensional covariance matrices", https://doi.org/10.1214/12-AOS989

Ledoit, O., Wolf, M. (2015), "Spectrum estimation: A unified framework for covariance matrix estimation and PCA in large dimensions", *Journal of Multivariate Analysis*, Vol. 139, No. 7, pp. 360-384. https://doi.org/10.1016/j.jmva.2015.04.006

Ledoit, O., Wolf, M. (2020), "Analytical non-linear shrinkage of large-dimensional covariance matrices", *The Annals of Statistics*, Vol. 48, No. 5, pp. 3043-3065. https://www.jstor.org/stable/27028732

Ledoit, O., Wolf, M. (2022), "The power of (non-)linear shrinking: A review and guide to covariance matrix estimation", *Journal of Financial Econometrics*, Vol. 20, No. 1, pp. 187-218. https://doi.org/ 10.1093/ jjfinec/nbaa007

Lutgens, F., Schotman, P. C. (2010), "Robust portfolio optimisation with multiple experts", *Review of Finance*, Vol. 14, No. 2, pp. 343-383. https://doi.org/10.1093/rof/rfn028

Nguyen, N., Nguyen, T., Tran, T., Mai, A. (2020), "Shrinkage model selection for portfolio optimization on

Vietnam stock market", *The Journal of Asian Finance, Economics and Business*, Vol. 7, No. 9, pp. 135-145. https://doi.org/10.13106/jafeb.2020.vol7.no9.135

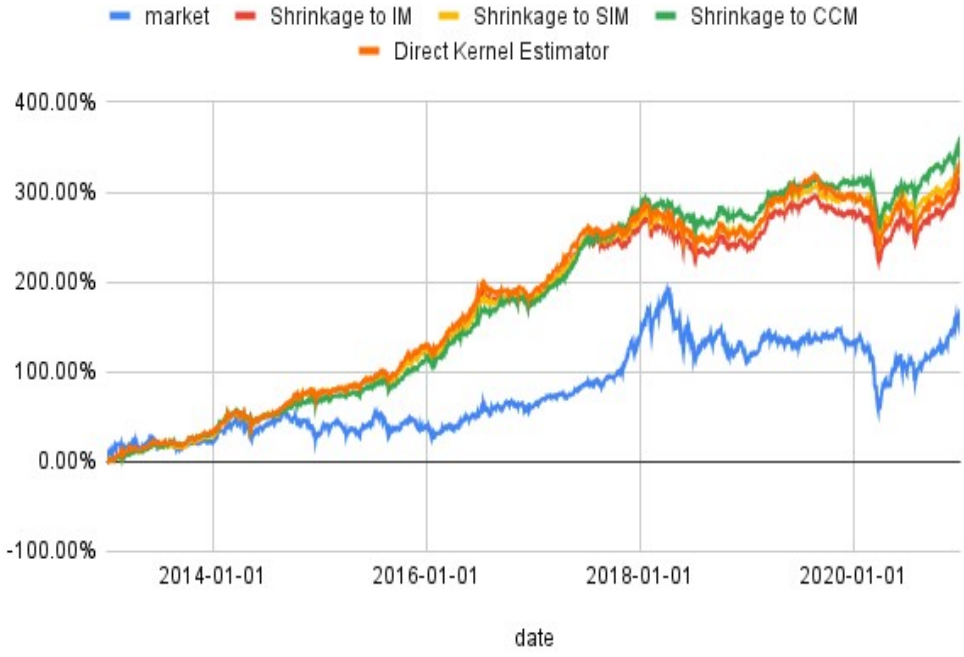
Schäfer, J., Strimmer, K. (2005), "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics", *Statistical Applications in Genetics and Molecular Biology*, Vol. 4, No. 1, pp. 1-15. https://doi.org/10.2202/1544-6115.1175

Scherer, B., Winston, K. (Eds.) (2011), "The Oxford handbook of quantitative asset management", OUP Oxford.

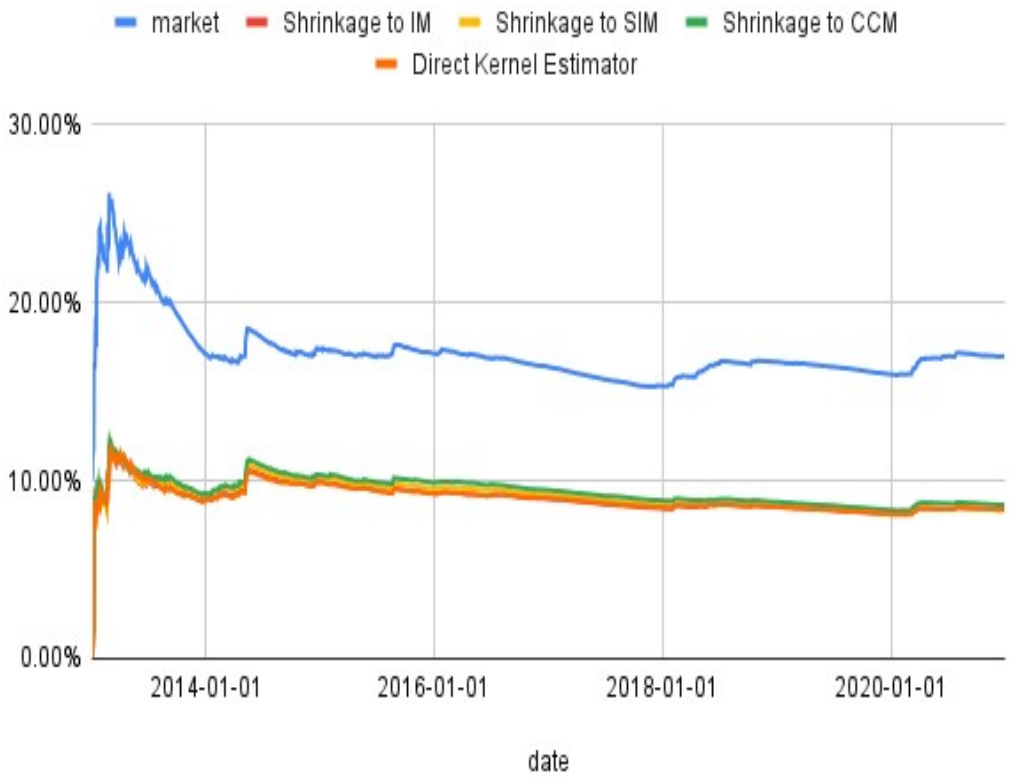
Tong, J., Hu, R., Xi, J., Xiao, Z., Guo, Q., Yu, Y. (2018), "Linear shrinkage estimation of covariance matrices using low-complexity cross-validation", *Signal Processing*, Vol. 148, No. 7, pp. 223-233. https://doi.org/10.1016/j.sigpro.2018.02.026

Tran, T., Nguyen, N., Nguyen, T., Mai, A. (2020, October), "Voting shrinkage algorithm for covariance matrix estimation and its application to portfolio selection", *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*, pp. 1-6. IEEE. https://doi.org/10.1109/RIVF48685.2020.9140764

Appendix A.



Cumulative returns of linear and non-linear shrinkage estimators on the whole Vietnamese stock market. The market line is the VNINDEX's return in the same period.



Volatility of linear and non-linear shrinkage estimators on the whole Vietnamese stock market. The market line is the VNINDEX's return in the same period.