UCLouvain

LOUVAIN School of Management

Louvain School of Management

How many assets to optimize out-of-sample performance?

Author: Louis Boufflette
Supervisor: Pr. Nathan Lassance

Academic year 2023-2024
Dissertation for the master 120 in Business Engineering
Financial Engineering
Daytime schedule

Abstract:

This master’s thesis explores theories and papers to find the optimal number of assets in one’s portfolio and conducts empirical research to find the optimal number of assets that maximize out-of-sample performance. The research will focus on specific portfolio optimization strategies and use two different methods of choice for the asset selection, one being random and the other using specific criteria.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management
Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique
www.uclouvain.be/lsm

Declaration Regarding AI Tool Usage in Master's Thesis

During the preparation of this master's thesis, the author(s) utilized Gifthub Copilot AI for the following purpose:

1.Reason: This AI was incorporated in the software R that I used to code my empirical research. This AI helped me code redundant task as repeating code and do the layout of some of my graphics.

2. After using Gifthub Copilot AI the author(s) diligently reviewed and edited the content produced by the tool. We take full responsibility for the final content presented in this thesis.

By signing this declaration, we affirm that the content of this master's thesis reflects our original work, augmented by the responsible use of AI.

Signed by Louis Boufflette on the date of the 20th May 2024.

Acknowledgments:

I would particularly like to thank my supervisor Professor Nathan Lassance, who took a lot of his time to help and guide me throughout the writing of my thesis. His expertise, availability and cooperation were the most valuable help I could have received, and I am very grateful for that.

I would also like to thank the entire teaching staff of my Financial Engineering major, especially Pr. Vrins whose course on derivative pricing helped me understand key concepts of finance and passed on to me his knowledge and passion for financial mathematics.

Last, I want to give a special thank my parents who supported me throughout all my university curriculum and made sacrifices to make my time as a student the best time of my life. Hopefully, I will be able one day to give it back to them. But I will certainly not stop thanking them for making me the man that I have become today.

Contents

1 Introduction

Since the last century, economists and investors have tried to determine the optimal portfolio of assets. Over the years, they have proposed and used various methods and models, from the famous modern portfolio theory of Markowitz (1952), which introduced the concept of portfolio diversification, to contemporary studies using advanced technologies and programming (Scaillet et al., 2023). There are several factors stemming from this topic that can be studied and discussed in detail such as weights, expected return, risk, and many more. But in the context of this master's thesis, we place the focus of this research on one important parameter: the number of assets within a portfolio of investment.

Many researchers have attempted to determine the optimal number of N assets in a portfolio. Zaimovic & al (2021) made a literature review on the topic and found various methods, outcomes, databases, and results aimed at identifying the optimal number of assets, starting with Evans & Archer (1968). Indeed, they examined the rate at which the variation of returns for randomly selected portfolios was reduced as a function of the number of assets included in the portfolio. They demonstrated that a relatively stable and predictable relationship exists between these variables, and throws considerable light on the decision as to the optimal number of securities to be included in a portfolio. Their results also raise doubts concerning the economic justification of increasing portfolio sizes beyond 10 or so securities. However, if we continue to explore the literature through time, the optimal number of assets fluctuates significantly to even reach more than 300 assets if we believe Haensly (2020), whose recent paper contributed to the literature on naïve diversification and risk decomposition. This difference of the optimal N number of assets in the literature illustrates the challenge of finding the optimal number of assets in a portfolio.

This challenge is the reason why I chose this topic. Consequently, this master's thesis aims to answer the question: "How many assets should one select to optimize out-of-sample performance?". Specifically, this master's thesis shows, in a particular setting, what the optimal number of assets is to achieve optimal portfolio diversification. We use theoretical knowledge from the literature while at the same time implementing an empirical analysis to test out-of-sample performance of different datasets using specific portfolio strategies. We will see whether it is efficient or not to invest in a large number of assets depending on the strategy you choose, knowing that controlling for the important issue of estimation risk still allows investors to profit from a non-negligible portfolio size. The underlying intuition is that increasing the number of assets provides more investment opportunities but also requires estimating additional weights, which introduces more estimation error. The optimal number of assets results from optimizing the trade-off between these two factors.

To address this, the thesis is divided into two main parts: a literature review/theoretical part and an empirical part.
In the first part, we start by doing a literature review using Zaimovic & al (2021). We introduce the review with Evans & Archer (1968) mentioned earlier and discuss the trend of the optimal number of assets (N) being between 10 and 20. We then develop the work of Statman (1987), who contested this trend by stating that an optimal portfolio should include at least 30 assets. We develop then the work of Lee & al. (2020) who discussed the viability of small account diversification for individual investors with a relatively smaller level of wealth. We conclude the review by discussing two

recent studies, Haensly (2020) and Arvanitis et al. (2023). Using different methods, these papers find that the optimal number of assets is 300 for the first and around 30 for the second. We then discuss our findings.

After the literature review, we transition to the theoretical part by presenting a review of the basic theoretical concept and methods regarding portfolio optimization using the pioneering paper of Markowitz (1952) from which we elaborate 3 critical assumptions regarding the investor's behaviour and their impact on the choice of N assets. We also elaborate the difference between in-sample and out-of-sample performance.

Then, we introduce the first case without parameter uncertainty and the concept of portfolio's utility. We show how it can be solved analytically and how utility evolves with the number of assets, N. In this case we see that the utility and the number of assets N have a positive linear relation meaning that the more assets we have, the better.

After that we present the case with parameter uncertainty, the latter being a more realistic representation of a real-life case. This is the reason why we need to explain how to estimate the mean and the covariance matrix of asset returns as it has a direct impact on the determination of the optimal N. We then show how N impacts our utility in this case. We observe that compared to the case without parameter uncertainty, the optimal N is small because it is directly linked with estimation errors and therefore worsens the performance of our portfolio when we add other assets in it.

To try to solve this problem, we present different robust methods that are used to improve performance when there is parameter uncertainty. We also explain what parameters are important to consider when talking about estimation errors. We present first two classic methods: the minimum-variance portfolio and the naive equally weighted portfolio and show how they perform compared to the sample mean-variance (SMV) portfolio. Before explaining what a shrinkage-based method is, we present the no-short-selling constraint which improves the performance of the SMV portfolio. After that we review a few shrinkage methods that we have seen in the literature such as Stein (1975) and dive deeper into the Ledoit & wolf shrinkage method for large-dimensional covariance matrices (2004).

Last, we show theoretically the optimal number of N assets for the two-fund rule of Kan & Zhou (2007), and demonstrate that N finds itself in the range $[0,T/2]$ which is bigger than when we do not control parameter uncertainty and naively invest in the SMV portfolio.

In our empirical part, we start in section 3.1 by presenting our 3 different datasets and their properties as they are used in the core of our analysis. We then dive into the details of our methodology.

Our research is divided into two main methods : The "Random" N method, where we use 100 random simulations of our datasets to build for each simulation 100 portfolios of size 1 to 100, and the " Smart" N method, where we use 3 different criteria to build for each criterion a set of 100 portfolios of size 1 to 100.

The criteria dictate how an investor would choose to add an asset to his portfolio. The criteria are : first, a choice based on the best in sample average return of an asset. Second, a choice based on the lowest variance and third a choice based on the momentum of an asset. The more an asset fit the criterion the sooner he is chosen to build our growing portfolio with the worst asset being the 100th to be chosen.

Regarding the choice of the optimal weights, we test these two methods on four portfolios: the

sample mean-variance, the minimum variance, the equally weighted, and the Kan and Zhou's two-fund rule portfolios.
To test out-of-sample performance, we evaluate the average out-of-sample utility using a rolling window technique, with the window size being equal to 120 months. This window is used to calculate optimal weights based on a defined portfolio strategy. Then, these same weights are multiplied by the returns of the 121st month to compute the out-sample-performance. This operation is repeated as the window moves in time to give us in total 507 different out-of-sample returns which we use to calculate the out-of-sample utility of the defined dataset and portfolio strategy.

Then, we compare the performances of the different portfolio strategies for both the random and smart methods and see whether we observe patterns of an optimal number (N) of assets in these portfolios. In section 3.3, we detail our results, where we observe that in every test, the sample mean-variance portfolio is always the worst performing portfolio. We also see that the best performing portfolio for the random method is the minimum variance which on counterpart has an irregular and unpredictable optimal number N of assets. For the strategies remaining, we have both equally weighted and Kan and Zhou portfolios that perform less than the minV, but have patterns regarding the optimal value of N being 10 to 15 and 75 to 95 respectively.

When comparing our smart criteria to our random results. We observe that the criteria improve the performance of every portfolio strategy except for the minimum variance. Furthermore, we see that the Kan and Zhou portfolio becomes the best performing portfolio for every dataset. We also have a decrease in the number of optimal N assets becoming 2 for the equally weighted and 15 to 35 for Kan and Zhou with the curve of the KZ strategy being close to the one seen in the theory developed in section 2.6.

To conclude this thesis, we wrap up every element that have been discussed before hypothesising about the limitation of this thesis and how it could be expanded or improved.

2 Literature Review & Theoretical Approach

In the following section we deep dive into the literature review and the theoretical approach of this thesis. First we start by doing a literature review on the problem of the optimal number of assets N. In 2.2, we develop the fundamental paper of Markowitz (1952) along with other basic concepts required to understand the rest of the thesis. Then, we define in 2.3 the mean-variance portfolio and develop its structure and also discuss the impact of N. In 2.4, we repeat the same structure but this time for the sample mean-variance portfolio (SMV) as it is exposed to parameter uncertainty. We explore different methods to alleviate this parameter uncertainty in 2.5 such as the equally weighted portfolio but also shrinkage-based methods such as the one of Ledoit & Wolf (2004). After that, we present in 2.6 the two-fund rule of Kan & Zhou (2007) and observe what the impact of N is if we use this method. Finally, we discuss in 2.7 some general perspective of our theoretical part.

2.1 Literature Review on the problem of optimal N assets

The research of the optimal number of assets N being our main focus for this literature review, we use different papers such as the one of Zaimovic & al. (2021) who analyze previous empirical studies to identify certain patterns in determining the optimal number of assets N in well-diversified portfolios in different markets and compare how the optimal number of assets N has changed over different periods. They say : *"Evaluating the number of N assets which lead to optimal diversification is not an easy task as it is impacted by a huge number of different factors: the way systematic risk is measured, the investment universe (size, asset classes and features of the asset classes), the investor's characteristics, the change over time of the asset features, the model adopted to measure diversification (i.e., equally weighted versus optimal allocation), the frequency of the data that is being used, together with the time horizon, conditions in the market that the study refers to, etc."*. The fact that this problem encompasses so many factors makes it such that our literature review has a large spectrum in terms of methods and findings. Still, we choose the following articles below because it offers a comprehensive representation of the literature in terms of diversity in approaches and methodologies, as well as the historical evolution of the problem.

2.1.1 Evans & Archer (1968)

One of the first paper we can find in the literature is the paper of Evans & Archer (1968). Indeed, one of the earliest and perhaps the most often quoted study of risk and portfolio size was performed by the two researchers, who developed a methodology that was to be used by almost all following studies. In their paper, They examine the rate at which the volatility for randomly selected portfolios on the US stock market is reduced as a function of the number of securities included in the portfolio. In doing so, they demonstrate that a relatively stable and predictable relation exists between these variables, and subsequently that the nature of this relation, if coupled with marginal cost considerations, throws considerable light on the decision as to the optimal number of securities to be included in a portfolio. Their results also raise doubts concerning the economic justification of increasing portfolio sizes beyond 10 or so assets, and indicate the need for analysts and private investors alike to include some form of marginal analysis in their portfolio selection models.

2.1.2 Statman (1987)

A trend that rises after the publication of the paper of Evans & Archer (1968) is that the diversification benefits measured by the volatility of portfolio returns are limited when we invest beyond 10 to 20 assets. Statman (1987) argues and wants to show that no fewer than 30 stocks are required for a well-diversified portfolio. He uses and expands the work of Elton & Gruber (1977) to show that a well-diversified portfolio must include at the very least 30 assets for a borrowing investor, and 40 assets for a lending investor. He also compares his result with the levels of diversification observed in studies of individual investors' portfolios and observes that people do not hold portfolios that are well diversified. Statman says this about the observation: *"Why do people forego the benefits of diversification? Maybe investors are simply ignorant about the benefits of diversification. If ignorance is the problem, education may be the solution. However, existing evidence does not warrant a claim that investors should indeed be educated to increase diversification".* This statement leads to numerous papers regarding the study of the behavioral portfolio theory. In his paper on the diversification puzzle, Statman (2004) observed that today's optimal level of diversification, measured by the rules of mean–variance portfolio theory, exceeds 300 assets, but the average investor holds only 3 or 4 assets. Indeed, to find this number Statman reminds that in the mean-variance portfolio theory, the optimal level of diversification is determined by marginal analysis meaning that diversification should be increased as long as its marginal benefits exceed its marginal costs. Using the original work of Markowitz (1952) and developing it to answer the question : " *what is the cost of the increased diversification to obtain that benefit?*, Statman (2004) computes that the risk reduction benefit of increasing diversification is optimal above 300 N assets.

2.1.3 Lee & al (2020)

It is in the context of the findings of Statman (2004) that Lee & al. (2020) discuss the viability of small account diversification for individual investors with a relatively smaller level of wealth than investors often classified as high-net-worth individuals. In order to reflect on the more practical situation for small investors, Lee & al. (2020) focus on the amount of investment for examining diversification benefits instead of the popular choice of using the number of securities in a portfolio. Furthermore, they analyze optimal mean-variance portfolios, rather than evaluating random portfolios to compute the average portfolio risk, where the optimal allocation is computed in terms of the number of assets N to incorporate the price of assets. Their results show that it is possible for individuals with low financial sustainability to invest in well-diversified portfolios, if not fully risk-diversified ones. They find out in the stock markets of six countries that an investment amount around 10,000 up to 100,000 US dollars is enough for constructing fully risk-diversified portfolios, which suggests that a substantial investment is not necessary for gaining complete diversification benefits. They also find that it is possible to form well-diversified portfolios with less than 10,000 US dollars, which is much less than the average amount invested in stocks in the U.S. Their findings show that forming a well-diversified portfolio within the mean-variance framework is achievable for average individuals. Finally, since the proposed quantity-based portfolio formulation with a budget constraint can be applied to any set of candidate assets, it can be used to form and analyze portfolios in certain sectors or industries. This may provide further details on how average individual investors can achieve risk diversification.

2.1.4 Haensly (2020)

In his recent paper, Haensly (2020) says that achieving adequate diversification is important in portfolio construction. He goes further by saying that : "*Efficient markets should not reward an investor for taking on risk that can be diversified away. Hence, when minimizing risk exposure, investors need to measure what part of total portfolio risk is systematic and what part can be diversified away.* Therefore, he examines several methods for decomposing total portfolio risk into systematic and diversifiable components and then carries out simulations to compare cross-sectional distributions of estimated and true risk as the number of stocks increases in portfolios constructed using naïve diversification. He observes that ordinary least squares estimators of diversifiable risk are relatively robust, and their cross-sectional distributions closely track the cross-sectional distributions of the corresponding true diversifiable risk. His results are relatively robust to the choice of method for generating market returns and to the underlying asset pricing model but not to random security betas. His simulation analysis also shows that risk and magnitude of shocks due to diversifiable risk are not negligible, even for portfolios with more than 300 assets.

2.1.5 Arvanitis & al (2023) and discussion

In their paper studying sparse spanning portfolios and under-diversification with second-order stochastic dominance, Arvanitis & al. (2023) develop and implement methods for determining whether relaxing sparsity constraints on portfolios improve the investment opportunity set for risk-averse investors. Using large equity datasets, they estimate the expected utility loss due to possible under-diversification, and find that there were no benefit from expanding a sparse opportunity set beyond 30 assets. The optimal sparse portfolio invests in 10 industry sectors with a larger weighting on small size, high book-to-market, and momentum stocks from the S&P 500 index and cuts tail risk when compared to a sparse mean-variance portfolio. On a rolling-window basis, the number of assets shrink to 10 assets in crisis periods.

To conclude this literature review, we could see that the optimal number N of assets varies for each paper that we present above. This shows that the statement of Zaimovic & al. (2021) exposed in 2.1 regarding the optimal number N of assets is in fact true. Let us keep that in mind for the next sections on the theoretical approach as it is discussed in section 2.7.

2.2 Problem of portfolio selection and foundational concepts

For the sake of this thesis, let's do a reminder of a few important parameters to pursue our research:

- r : The vector of the excess returns of the N assets
- μ : The mean of the returns
- Σ : The covariance matrix of the returns which is always assumed positive-definite
- w: The vector of the weights of the portfolio

We can denote the portfolio return as $r_P = w^\top r$ with expected return μ and variance $w^\top \Sigma w$.

Throughout this thesis, we have to consider two assumptions given by Markowitz (1952). The first assumption supposes that investors always tend to maximise their expected return, reflecting a fundamental principle in financial decision-making. Investors are presumed to be rational actors who, when faced with various investment choices, consistently try to optimize their portfolio's performance by seeking the highest possible expected return. The second assumption extends the first one by stating that investors consider expected return as a desirable thing and variance of return or volatility as an undesirable thing. This perspective gives us the intuition that investors weigh both potential gains and losses when making decisions, seeking a balance between maximizing returns and minimizing the perceived risk associated with fluctuations in investment performance. This phenomenon is what we refer as the mean-variance trade-off which investors try to maximise always taking into account that the higher the risk, the higher the potential return. This can be graphically represented in Figure 1 (Hull, 2018), where you can see the linear relation between risk and return.

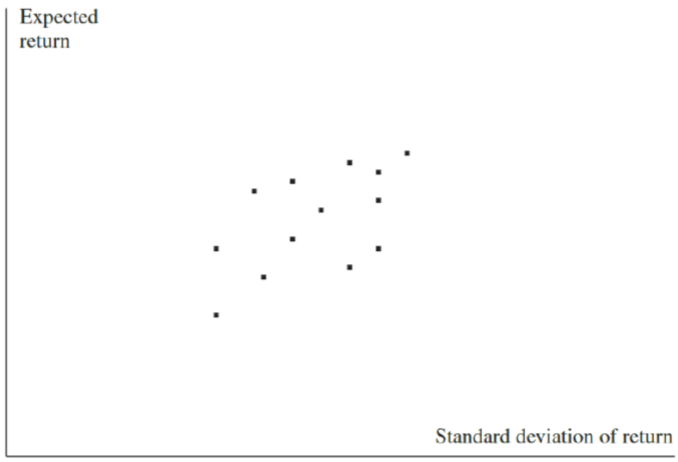

Figure 1: Linear relation between risk and return (Hull, 2018)

As the focus of this thesis is on the number N of assets, we can discuss its potential impact on the two previous assumptions, as we want to find the optimal N rather than it being fixed. Indeed as N increases, investors may have more opportunities to diversify their portfolios. Diversification involves investing in a variety of assets to spread risk. With a larger number of assets N, investors could potentially construct portfolios that offer a more balanced risk-return profile. A large number of assets allows investors to also consider not only the returns of individual assets but also the correlations among them. The impact of N on the mean variance trade-off is tied to the diversification effect. With a well-diversified portfolio (larger N), investors may be able to achieve a certain level of expected return with lower volatility or variance. This emphasizes the importance of diversification in managing risk.
A third critical assumption is that theoretical frameworks often assume that investors making optimal financial decisions know the true parameters of the model indicating that there is no parameter uncertainty. This implies that investors are presumed to have precise information about key parameters such as the mean μ or the covariance matrix Σ. Based on all that information, we present in the next section the mathematical solution of the mean-variance portfolio and show how the number N of assets influence its performance.

2.3 Mean-Variance portfolio

In Markowitz's mean-variance portfolio theory (1952), we consider here the case where the universe includes a risk-free asset, meaning that the portfolio is composed of N risky assets and one risk-free asset. This is represented by an efficient frontier that can be graphically represented in Figure 2 (Hull, 2018) where all the portfolios on this frontier are called efficient mean-variance portfolios

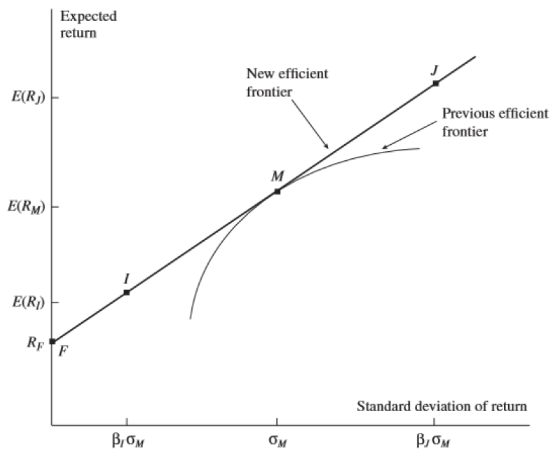

Figure 2: Efficient frontier of Markowitz (Hull, 2018)

In this section, we assume that the third assumption of the section 2.2 holds true, meaning that there is no parameter uncertainty. Markowitz also introduced the concept of utility as a measure of an investor's satisfaction or well-being from a particular portfolio. He argued that investors make decision not only based on expected return and risk but also take into account their personal preferences and attitudes towards risk. This notion allows for a more realistic representation of investor decision making by incorporating individual preferences and risk tolerance into the portfolio selection process.

A mean-variance investor allocates wealth according to the portfolio w that solves the following problem :

$$w^* = \underset{w \in \mathbb{R}^N}{\operatorname{argmax}}\, U(w), \tag{1}$$

where $U(w)$ is the utility of the portfolio w,

$$U(w) = w^\top \mu - \frac{\gamma}{2} w^\top \Sigma w, \tag{2}$$

and $\gamma > 0$ is the investor's risk aversion coefficient. The solution to (1) is

$$w^* = \frac{1}{\gamma} \Sigma^{-1} \mu, \tag{3}$$

and the weight on the risk-free asset is $1 - 1_N^\top w^*$, where 1_N is a $N \times 1$ vector of ones. The maximum utility delivered by w^* is

$$U(w^*) = \frac{\theta^2}{2\gamma} \tag{4}$$

where

$$\theta^2 = \mu^\top \Sigma^{-1} \mu \tag{5}$$

is the maximum squared Sharpe ratio (Sharpe, 1966), a risk-adjusted performance measure for a portfolio. In the context of finding an optimal N for a portfolio, we can ask ourselves : which N maximises $U(w^*)$? At first thought, we can have the intuition that the larger N the better, as it yields more investment opportunities. To formalize this, let us assume that all correlations are equal, i.e.,

$$\Sigma = DCD, \tag{6}$$

where D = diag(σ_1,...,σ_N) is a diagonal matrix containing the volatilities, and the correlation matrix C is of the form

$$C = (1-\rho)I_N + \rho 1_N 1_N^\top, \tag{7}$$

where ρ is the correlation across all pairs of assets. The inverse of Σ in (6) is

$$\Sigma^{-1} = D^{-1}C^{-1}D^{-1}, \tag{8}$$

Where $D^{-1} = \text{diag}(\frac{1}{\sigma_1}, ..., \frac{1}{\sigma_N})$ and C^{-1} exists if $-\frac{1}{N-1} < \rho < 1$ and is equal to

$$C^{-1} = \frac{1}{1-\rho}I_N - \frac{\rho}{(1-\rho)(1+(N-1)\rho)}1_N 1_N^\top. \tag{9}$$

If we combine the equations (8) and (9) together we obtain

$$\Sigma^{-1} = \frac{1}{1-\rho}D^{-2} - \frac{\rho}{(1-\rho)(1+(N-1)\rho}D^{-1}1_N 1_N^\top D^{-1}, \tag{10}$$

and under the specific form of Σ^{-1} in equation (10), the maximum utility from (4) becomes

$$U(w^*) = \frac{\mu^\top \Sigma^{-1}\mu}{2\gamma} = \frac{1}{2\gamma(1-\rho)}[\mu^\top D^{-2}\mu - \frac{\rho}{1+(N-1)\rho}(\mu^\top D^{-1}1_N)^2]. \tag{11}$$

Let us denote $\theta_i = \frac{\mu_i}{\sigma_i}$, the Sharpe ratio of asset i. Then, the quantities $\mu^\top D^{-2}\mu$ and $\mu^\top D^{-1}1_N$ in (11) depend on the θ_i's as

$$\mu^\top D^{-2}\mu = \sum_{i=1}^{N}\theta_i^2, \tag{12}$$

$$\mu^\top D^{-1}1_N = \sum_{i=1}^{N}\theta_i. \tag{13}$$

Denoting $\bar{\theta}_1 = \frac{1}{N}\sum_{i=1}^{N}\theta_i$ and $\bar{\theta}_2 = \frac{1}{N}\sum_{i=1}^{N}\theta_i^2$, we finally have that $U(w^*)$ in (11) becomes

$$U(w^*) = \frac{\theta^2}{\gamma} \quad \text{with} \quad \theta^2 = \frac{N}{1-\rho}\left[\bar{\theta}_2 - \frac{\rho N}{1+(N-1)\rho}\bar{\theta}_1^2\right]. \tag{14}$$

Note that we can show that $\bar{\theta}_2 \geq \bar{\theta}_1^2$. This formula shows that $U(w^*)$ is an increasing function of N and actually grows to infinity as $N \to \infty$. This suggests investing in as many stocks as possible. There is also a dependency on N in $\bar{\theta}_1$ and $\bar{\theta}_2$, but it fades out as N becomes large. In Figure 3 below, we assume a monthly return frequency and we set $\gamma = 1$, $\bar{\theta}_1 = 0.15$, $\bar{\theta}_2 = \bar{\theta}_1^2 + 0.01$, and we depict how $U(w^*)$ depends on N for $\rho = 0.2$, 0.5 and 0.8.

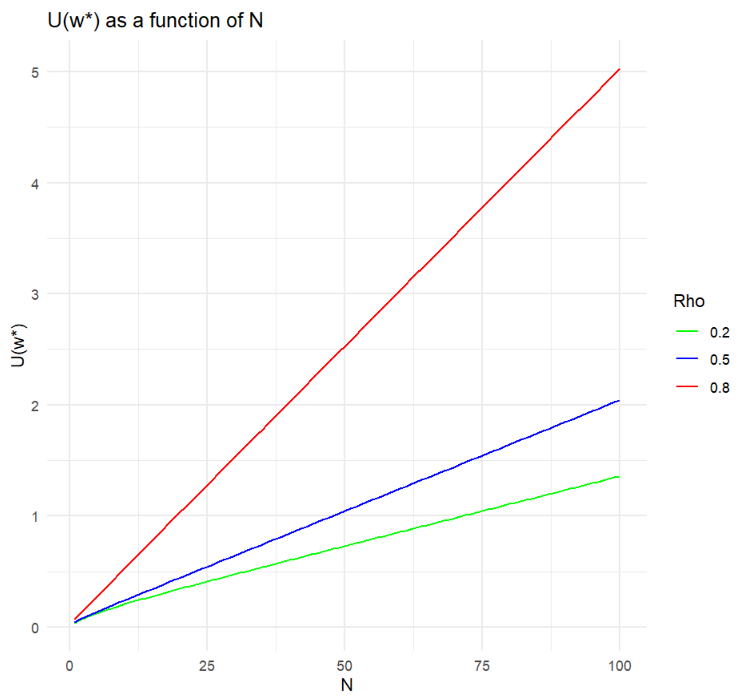

Figure 3: $U(w^*)$ as a function of N and ρ

2.4 Sample mean-variance portfolio

In the previous section, we saw that if investors knew the true parameters of the model, then the utility function would be an increasing function of N. But the problem with this assumption is that the true parameters are rarely, if ever, known to decision makers (Kan & Zhou, 2007). The previous analysis is then not feasible in practice because the portfolio w^* depends on the true μ and Σ, which are actually unknown. Instead, parameters of the model such as μ and Σ have to be estimated. In practice, we collect $T > N$ vectors of historical stocks returns, $(r_1,.....r_T)$, and the sample estimators of μ and Σ are

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_t, \tag{15}$$

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}(r_t - \hat{\mu})(r_t - \hat{\mu})^\top. \tag{16}$$

As a reminder, the requirement $T > N$ is driven by the need for a sufficient number of data points to reliably estimate the parameters of the model as it reduces overfitting and ensure that the covariance matrix remains invertible. Given a portfolio $\hat{w}$ constructed at time T from $\hat{\mu}$ and $\hat{\Sigma}$, its out-of-sample utility at time $T+1$ is

$$U(\hat{w}) = \hat{w}^\top \mu - \frac{\gamma}{2}\hat{w}^\top \Sigma \hat{w}, \tag{17}$$

One portfolio we might then think to consider is the plug-in estimator of w^* ,i.e.,

$$\hat{w}^* = \frac{1}{\gamma}\hat{\Sigma}^{-1}\hat{\mu} \tag{18}$$

However, w^* performs notoriously poorly out-of-sample. Indeed , the main difficulty in estimating the mean-variance portfolio resides in the estimation of the mean returns μ for two principal reasons:

- 1.The variance of the sample $\hat{\mu}$ is generally much higher than the variance of the sample variances and co-variances.

Black (1993), says that estimating the expected return poses a significant challenge, with daily data providing limited assistance. Accurate estimates of the average expected return require an extensive dataset spanning decades and it's crucial to acknowledge the portfolio's high sensitivity to this estimation.

- 2. The mean variance portfolio is very sensitive to the sample mean $\hat{\mu}$

Best and Grauer (1991) also show that the SMV portfolio displays significant instability, mainly attributed to the heightened sensitivity of the optimal solution to minor adjustments in input assumptions. Even slight changes in these assumptions can induce substantial divergence, leading to totally different optimal outcomes. Therefore, the SMV portfolio $\mathbb{E}[\hat{w}^*] = \frac{T}{T-N-2}w^*$ is very sensitive to estimation error in the inputs which has two severe impact:

- It deteriorates its out-of-sample performance
- It is very unstable over time , which requires intensive re-balancing and high transaction costs which is something we don't want in the context of this work

What about our number of N assets ? To understand its impact on the out-of-sample performance of the SMV portfolio, let's come back to equation (17) where we want to find $\hat{w}^*$ maximising the expected out-of-sample utility, $\mathbb{E}[U(\hat{w})]$. By doing some computations and assuming that $r_t \sim \mathcal{N}(\mu, \Sigma)$ and that all T observations are i.i.d, provided $T > N + 4$ we find out the following expression (Kan and Zhou, 2007) :

$$\mathbb{E}[U(\hat{w})] = \frac{1}{\gamma}\frac{T\theta^2}{T-N-2} - \frac{1}{2\gamma}\frac{(T-2)(T\theta^2+N)}{(T-N-1)(T-N-2)(T-N-4)}. \tag{19}$$

If we compare this expression to the one without parameter uncertainty, where $U(w^*) = \frac{\theta^2}{2\gamma}$ is the maximum without estimation risk, we can see that our equation it is strictly smaller $U(w^*)$ because of parameter uncertainty. But unlike w^*, $\hat{w}^*$ is actually feasible. Which N maximise $\mathbb{E}[U(\hat{w})]$ then ? Let's incorporate θ^2 in our previous equation. Doing so we find :

$$\mathbb{E}[U(\hat{w})] = \frac{T(\frac{N}{1-\rho}[\bar{\theta}_2 - \frac{\rho N}{1+(N-1)\rho}\bar{\theta}_1^2])}{\gamma(T-N-2)} - \frac{(T-2)(T(\frac{N}{1-\rho}[\bar{\theta}_2 - \frac{\rho N}{1+(N-1)\rho}\bar{\theta}_1^2]) + N)}{2\gamma(T-N-1)(T-N-2)(T-N-4)}. \tag{20}$$

Unfortunately there is no easy expression to explicitly represent the impact of N on the equation. But, we can still use this equation to graphically represent its impact. In Figure 4, we assume a monthly return frequency and we set $\gamma = 1$, $T = 120$, $\bar{\theta}_1 = 0.15$, $\bar{\theta}_2 = \bar{\theta}_1^2 + 0.01$, and we depict how $\mathbb{E}[U(\hat{w}(c^*))]$ depends on N for $\rho = (0.2, 0.5, 0.8)$. We directly observe that the function is a decreasing function of N and ρ. Confirming our intuition that in the case of the SMV portfolio, the optimal N is very small as increasing its value leads to a too large increase in estimations error.

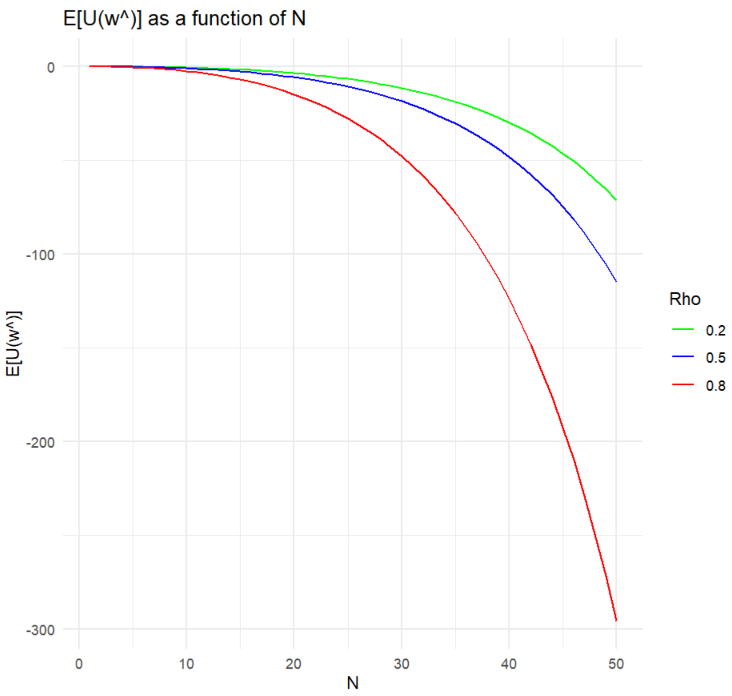

Figure 4: $\mathbb{E}[U(\hat{w})]$ as a function of N and ρ

If we zoom in Figure 4, we can observe that the optimal N is between 1 and 2 as we can see on Figure 5 below.

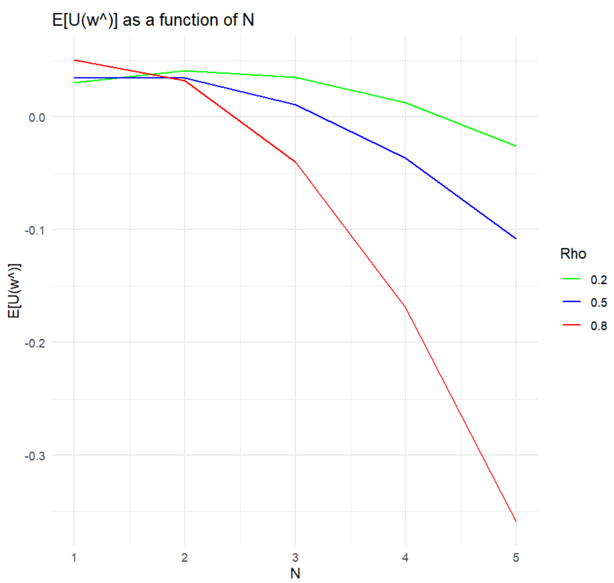

Figure 5: A zoom on $\mathbb{E}[U(\hat{w})]$ as a function of N and ρ

What should we do then to improve the estimation of the mean-variance portfolio and by extension allows for a larger number N of assets? We explore different methods and solutions in the next sections.

2.5 Alleviating parameter uncertainty

We have seen that the optimal number of assets N for the SMV portfolio is generally small because of the lack of robustness to parameter uncertainty leading to an increasing estimation risk when we add assets in the portfolio. Indeed, Michaud (1989) argues that optimal SMV portfolios are often prone to "error maximization".

We also know that there are many different factors that determine how robust your estimated portfolio should be.

- Stability of the asset return distribution: the less it is stable , the more robust estimation should be
- Sample size T: the lower the sample size, the more robust your estimation should be
- Number of assets N: the higher the number of assets, the more robust your estimation should be

With the purpose of reducing the effect of estimation errors in the estimates of mean values and variances, more stable MV portfolio selection models are formulated by applying robust techniques. We present in this section a non-exhaustive list of robust methods well settled in the scientific landscape to alleviate the bad performance of the SMV portfolio. Methods can focus on different parameters, the estimation of $\hat{\mu}$ that we mentioned in the previous section, but also on shrinkage methods of $\hat{\Sigma}$ and the weights $\hat{w}$.

2.5.1 Minimum-variance portfolio

Regarding $\hat{\mu}$, one possibility is to improve the estimation of the latter would be to simply assume that the sample mean is the same for all assets, i.e., $\mu = c1_N$, in which case the SMV portfolio corresponds to the minimum-variance portfolio whose weights are given by the following :

$$\frac{\hat{\Sigma}^{-1} 1_N}{1_N^\top \hat{\Sigma}^{-1} 1_N}. \tag{21}$$

But although, the minimum-variance portfolio is considered to be more robust than the mean-variance portfolio and has better performance out-of-sample (DeMiguel & Nogales, 2007), this method remains sensitive to estimation error in the covariance matrix $\hat{\Sigma}$, especially when N is increasing towards T.

2.5.2 Equally weighted portfolio

But what if the best way to reduce estimation error was actually not to estimate anything at all? This naïve diversification strategy is what we call the equally weighted portfolio in which investors invest the same weights in all assets. The solution to this basic strategy is then given by :

$$w_{EW} = \frac{1_N}{N}. \tag{22}$$

Though this strategy seems very naïve, DeMiguel & al. (2009) observed that, in real-world settings where estimation error is often present, this approach offers a good performance compared to the

SMV portfolio but also more complex portfolio optimizations. Indeed, DeMiguel & al. (2009) reported "*we find that of the various optimizing models in the literature, there is no single model that consistently delivers a Sharpe ratio that is higher than that of the 1/N portfolio, which also has a very low turnover*". However, this strategy can also have some drawbacks, especially when assets have significantly different individual risks and/or significantly different correlations. In the end, the 1/N naïve diversification rule should serve at least as an obvious benchmark for portfolio performance.

2.5.3 No-short-selling constraint

To reduce estimation error, we can try to apply a short-sale constraint on our SMV portfolio. We can obtain it by imposing an additional non-negativity constraint on the portfolio weights in the optimization problem. DeMiguel & al. (2009) developed the interpretation of the effect of short-sale constraints to state that imposing a short sale constraint on the SMV portfolio is equivalent to "shrinking" the expected return towards the average. Jagannathan & Ma (2003) also showed that imposing a short-sale constraint on the minimum-variance portfolio presented in subsection 2.5.1 is equivalent to "shrinking" the elements of the covariance matrix. They say "*the sample covariance matrix performs almost as well as those constructed using factor models, shrinkage estimators or daily returns.*" Even though this constraint improves the result of the out-of-sample performance of the SMV portfolio, it still were worse than the equally weighted portfolio as shown by DeMiguel & al. (2009).

2.5.4 Shrinkage-based methods

As we explained in the beginning of this section, the covariance matrix of the asset returns is key component to explain estimation errors. In real datasets, when the number N of assets is large relative to the available number T of historical return observations, we observe an important number of estimation errors in the sample covariance matrix leading to large errors on the optimal weights of our portfolio. To avoid those estimation errors in the covariance matrix, Stein (1956) introduced the statistical concept of shrinkage. The objective of shrinking the sample covariance matrix is to pull the most extreme values towards a more central ones so that it reduces the impact of estimation errors when needed. The problem is to know what the optimal shrinkage of the sample covariance matrix is. Many papers in the literature work on this challenge. From the Stein's shrinkage estimator of eigenvalues (1975) whose method is based on the observation that unless $T >> N$, the eigenvalues of the sample covariance matrix $\hat{\Sigma}$ are severely distorted compared to those of the true covariance matrix Σ, to shrinkage methods based on the combination of different portfolios such as Tu & Zhou (2011) who introduced a combination of the Markowitz method with the naive $1/N$ portfolio according to a combination coefficient δ. There are still many areas to cover but for the sake of this thesis, we only elaborate the one of Ledoit & Wolf (2004) as the last part of this subsection.

2.5.5 Ledoit & Wolf

The shrinkage of the covariance matrix with the estimator of Ledoit & Wolf (2004b) aims as every shrinkage method at reducing its volatility and improving its stability especially when T is limited compared to N. This method is considered to be more stable and robust and is given by the following equation:

$$\hat{\Sigma}_{LW} = (1-\delta)\hat{\Sigma} + \delta\hat{T}, \tag{23}$$

where $\hat{\Sigma}_{LW}$ is a linear combination of the unbiased-high-variance sample estimate $\hat{\Sigma}$ and a biased-low-variance target estimate $\hat{T}$. For the target estimate, one choice is to assume that the assets returns are uncorrelated and have the same variance : $\hat{T} = \sigma^2 I_n$ (Ledoit & Wolf, 2004) and where $\delta \in$[0,1] is the shrinkage intensity computed as an estimate of the minimizer of the Frobenius norm of $\hat{\Sigma}_{LW} - \Sigma$:

$$\delta = \underset{\delta\in[0,1]}{\operatorname{argmin}} ||\hat{\Sigma}_{LW} - \Sigma||_F, \tag{24}$$

where

$$||A||_F = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}^2}. \tag{25}$$

In essence, Ledoit & Wolf (2004) propose a method that involves reducing the size of the sample covariance matrix, guided by a specified target matrix $\hat{T}$. This target matrix is designed to encapsulate crucial characteristics of the estimated quantity while incorporating only a limited number of parameters, resulting in a highly structured representation. The adjusted matrix, derived through this process, can then be effectively employed, for example, in the calculation of weights for constructing a minimum variance portfolio :

$$\hat{w}_{MSR,LW} = \frac{\hat{\Sigma}_{LW}^{-1}1}{1^{\top}\hat{\Sigma}_{LW}^{-1}1} \tag{26}$$

2.6 Kan & Zhou two-fund rule

In the continuation of the last section regarding alleviating the parameter uncertainty, we study in this section another shrinkage method, introduced by Kan & Zhou (2007), that reduces the impact of estimation risk by optimally combining the SMV portfolio with the risk-free asset so that we can maximise the out-of-sample performance. This methodology is not unique , we can find in the literature other papers where the authors worked on the shrinkage of weights by combining different portfolios strategies. We already mentioned Tu & Zhou (2011) in section 2.5 who used combinations of the $1/N$ naive portfolio, but there is also Lassance & al. (2024) who showed that neither the SMV portfolio nor the sample minimum-variance portfolio exhibited maximal robustness individually, and that investors needed to combine both to optimize portfolio robustness. Kan & al (2022) also discussed the combining strategy of the SMV without the risk-free asset providing not only a fast and accurate way of evaluating the performance, but also analytical insights into the portfolio construction.

The reason why we chose to study the paper of Kan & Zhou (2007) is that this particular approach makes it theoretically possible to find the corresponding optimal N number of assets which is what we are aiming at. Kan & Zhou (2007) propose a two-fund rule that combines w^* and the risk-free asset:

$$\hat{w}(c) = c\hat{w}^* = \frac{c}{\gamma}\hat{\Sigma}^{-1}\hat{\mu}, \tag{27}$$

where c $\in \mathbb{R}$ is a combination coefficient to determine. Given that $U(\hat{w}(c))$ is a random variable, Kan & Zhou (2007) propose to find the value of c maximizing the expected out-of-sample utility :

$$c^* = \underset{c\in\mathbb{R}}{\operatorname{argmax}} \ \mathbb{E}[U(\hat{w}(c))] = \underset{c\in\mathbb{R}}{\operatorname{argmax}} \ \mathbb{E}[\hat{w}(c)^\top \mu - \frac{\gamma}{2}\hat{w}(c)^\top \Sigma \hat{w}(c)] \tag{28}$$

Assuming $r_t \sim \mathcal{N}(\mu, \Sigma)$ and that all T observations are i.i.d, Kan & Zhou (2007) show that, provided $T > N + 4$,

$$\mathbb{E}[U(\hat{w}(c))] = \frac{c}{\gamma}\frac{T\theta^2}{T-N-2} - \frac{c^2}{2\gamma}\frac{(T-2)(T\theta^2+N)}{(T-N-1)(T-N-2)(T-N-4)}, \tag{29}$$

which yields

$$c^* = \frac{(T-N-1)(T-N-4)}{T(T-2)}\frac{\theta^2}{\theta^2+\frac{N}{T}}. \tag{30}$$

Pay attention that $0 < c^* < 1$, i.e., it is optimal to tone down the exposure to the SMV portfolio $\hat{w}^*$, but not completely. Plugging c^* into (29), the expected out-of-sample utility of the optimal two-fund rule $\hat{w}(c^*)$ is

$$\mathbb{E}[U(\hat{w}(c^*))] = U(w^*)\frac{(T-N-1)(T-N-4)}{(T-2)(T-N-2)}\frac{\theta^2}{\theta^2+\frac{N}{T}}, \tag{31}$$

which is strictly smaller than $U(w^*)$ in (4) because of parameter uncertainty. Unlike w^* that is unfeasible, $\hat{w}(c^*)$ is feasible, with the only caveat that c^* is unknown in practice because it depends on the true θ^2.

Again we can ask ourselves the question: which N maximizes $\mathbb{E}[U(\hat{w}(c^*))]$? On one hand, we saw in the setting without parameter uncertainty that θ^2 increases with N. On the other hand, the quantity $\frac{(T-N-1)(T-N-4)}{(T-2)(T-N-2)}$ decreases with N, and $\frac{N}{T}$ in the denominator of (30) also goes against increasing N. To formalise this trade-off, let us replace θ^2 by its simplified formula in (14). This gives

$$\mathbb{E}[U(\hat{w}(c^*))] = \frac{1}{2\gamma}\frac{(T-N-1)(T-N-4)}{(T-2)(T-N-2)}\frac{(\frac{N}{1-\rho}[\bar{\theta}_2 - \frac{\rho N}{1+(N-1)\rho}\bar{\theta}_1^2])^2}{\frac{N}{1-\rho}[\bar{\theta}_2 - \frac{\rho N}{1+(N-1)\rho}\bar{\theta}_1^2] + \frac{N}{T}} \tag{32}$$

which we can transform into

$$\frac{N}{2\gamma(1-\rho)}\frac{(T-N-1)(T-N-4)}{(T-2)(T-N-2)}\frac{(\bar{\theta}_2 - \frac{\rho N}{1+(N-1)\rho}\bar{\theta}_1^2)^2}{\bar{\theta}_2 - \frac{\rho N}{1+(N-1)\rho}\bar{\theta}_1^2 + \frac{1-\rho}{T}} \tag{33}$$

This is now an explicit function of N. Unfortunately, there is no easy expression for the optimal value of N. In Figure 6, we assume a monthly return frequency and we set $\gamma = 1$, $T = 120$, $\bar{\theta}_1 = 0.15$, $\bar{\theta}_2 = \bar{\theta}_1^2 + 0.01$, and we depict how $\mathbb{E}[U(\hat{w}(c^*))]$ depends on N for $\rho = (0.2, 0.5, 0.8)$. Figure 6 shows that the optimal value of N is slightly smaller than $T/2 = 60$. To explain why this is the case, note that $\frac{(T-N-1)(T-N-4)}{(T-2)(T-N-2)} \approx 1 - \frac{N}{T}$, and thus we can approximate $\mathbb{E}[U(\hat{w}(c^*))]$ as a product of two terms :

$$\mathbb{E}[U(\hat{w}(c^*))] \approx \frac{N(1-\frac{N}{T})}{2\gamma(1-\rho)} \times \frac{(\bar{\theta}_2 - \frac{\rho N}{1+(N-1)\rho}\bar{\theta}_1^2)^2}{\bar{\theta}_2 - \frac{\rho N}{1+(N-1)\rho}\bar{\theta}_1^2 + \frac{1-\rho}{T}} \tag{34}$$

The first term is maximised by $N = T/2$ while the second one decreases as N increases, which explains why the optimal N is be between zero and $T/2$.

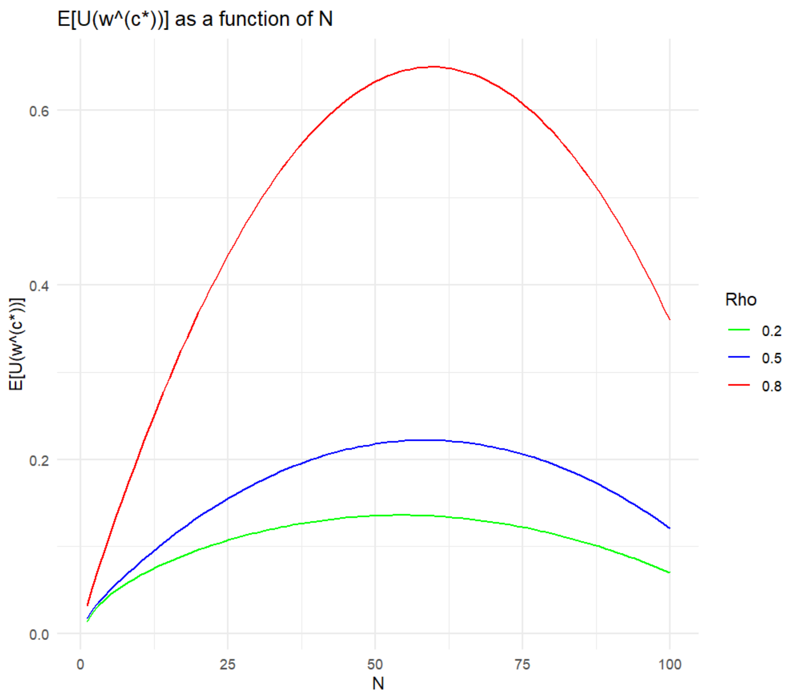

Figure 6: $\mathbb{E}[U(\hat{w}(c^*))]$ as a function of N and ρ

2.7 General perspective

As we said at the end of our literature review in section 2.1 when we discussed about the statement of Zaimovic & al. (2021) regarding the evaluation of the optimal N numbers of assets, we can have the intuition that the optimal value of N is intimately tied to the portfolio strategy we consider. Each different portfolio strategy we discussed or analyzed yielded a different optimal N. In particular, the more the strategy is affected by parameter uncertainty, the lower the optimal N. If we compare the SMV versus the two-fund rule of Kan & Zhou, we see that the optimal N in the SMV is lower than in the other as it is more exposed to estimation risks.
If we take the example of the last strategy analyzed (Kan & Zhou), the theoretical analysis relies on a number of assumptions, in particular, asset returns are i.i.d. normal and we know the optimal c^*. These assumptions can be discussed, if we take the paper of Kan & Lassance (2024), the authors showed that the assumption of normality for asset returns was not innocuous because fat tails in returns increase the out-of-sample mean and variance of sample portfolios relative to normality showing that assumptions can influence the result of one strategy. Moreover, more sophisticated strategies may not be possible to analyze theoretically. Thus, this theoretical analysis must be complemented with empirical evidence in which we assess how the out-of-sample performance delivered by different portfolio strategies depends on N in different datasets. This is what we do in the next section of this thesis.

3 Empirical approach

This part of the thesis is dedicated to our empirical research. First, we start in 3.1 by presenting the different datasets that we use throughout our research with explanations and assumptions regarding their use. In the next section 3.2, we present the methodology of tests and research we made in the Software R. We introduce our two main different methods being ”Random” N and ”Smart” N and afterwards we present the different portfolios strategies we choose to conduct our experiment. We discuss the rolling window technique we use and also on certain key parameters and the choices we make regarding them. Section 3.3 regroups our results and analysis.

3.1 Datasets

To conduct our analysis on the optimal number of N assets on the different evaluated portfolio strategies, we use throughout all this part 3 different datasets which were downloaded from the Kenneth French Data Library :

1. The first dataset is from the *100 Portfolios Formed on Size and Book-to-Market (10 x 10)* which is constructed in a way such that portfolios are formed at the end of each June and result from the intersection of 10 portfolios based on size (market equity, ME) and 10 portfolios based on the ratio of book equity to market equity (BE/ME). The size (ME) is the size breakpoints for year t are the NYSE market equity deciles at the end of June of year while the BE/ME ratio for June of year t is calculated by taking the book equity for the last fiscal year ending in t-1 and dividing it by ME for December of t-1. The BE/ME breakpoints are NYSE deciles. The stocks included in the portfolios for July of year t to June of year t+1 include all NYSE, AMEX, and NASDAQ stocks for which there is market equity data for December of t-1 and June of t, as well as book equity data for t-1. The dataset will be called ”SBM” for the rest of the thesis for the sake of clarification.

2. The second dataset is from the *100 Portfolios Formed on Size and Operating Profitability.* The difference with the first one is that here we use Operating Profitability (OP) meaning that portfolios are formed based on operating profitability, calculated as annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses, divided by book equity for the last fiscal year end in t-1. The dataset will be called ”SOP” for the rest of the thesis.

3. The Last dataset is from the *100 Portfolios Formed on Size and Investment.* Here we use investment implying that portfolios are formed based on investment, calculated as the change in total assets from the fiscal year ending in year t-2 to the fiscal year ending in t-1, divided by t-2 total assets. This dataset will be called ”SINV” for the rest of this thesis.

The time period for the datasets is from January 1970 to March 2022 and is on a monthly frequency meaning that we have in total 627 historical observations for each. In the datasets, missing data are represented at a Value of -99.99% which is not desirable. Therefore, we take the decision to replace the missing values by 0 as this method is common in the literature. Other methods could have been implemented such as replacing by the sample mean of the variable or by the sample mean of the entire dataset but the fact that the number of missing value is very small makes it such that the results are intuitively not sensitive to the way they are replaced.

To transform the returns of our datasets into excess returns, we also downloaded the monthly risk free rate on the Kenneth French Library from the Fama/French 3 factors which we subtracted from our datasets' returns. Using these datasets allows us to investigate on the optimal number of N assets in different scenarios due the difference within the datasets. More information about the datasets can be found on the Kenneth R.French website.

3.2 Methodology

Now that we have presented our datasets, we can discuss how we are going to manipulate them. As briefly mentioned in the intro of this part, we decided to select two different approaches to conduct our analysis regarding the optimal choice of the parameter N. We call the first one "Random" N while the other will be called "Smart" N.

3.2.1 Random N

The "Random" N method goes as follows : among our 3 different datasets (SBM, SOP and SINV) we have 100 different assets. "Random" N means that from our original dataset of 100 assets, we build growing portfolios whose assets are randomly chosen. We specify that if an asset is randomly chosen for N =1, this same asset will be in the portfolio $N = 2$. The outcome of this randomisation is then 100 portfolios with the first portfolio having one asset and the 100th having the same composition as our original dataset.
What matters in this randomisation is to see whether the random path used to make our portfolio grown in size has an impact on the optimization of N. As one simulation would be arbitrary and lead to a nonobjective representation, we decided to repeat this randomisation multiple times. We chose to repeat a hundred times the randomisation giving us better output to work with. The choice of hundred is arbitrary but also due to a constraint linked to our limited computational power.

3.2.2 Smart N

The random N method is limited by its very nature. In a real-life situation, this means that an investor would randomly select what he invests in which is unlikely unless the investor does not have any information on the dataset or does not know how to select assets. Intuitively, we can think that an investor tries to choose what he invests in based on certain criteria. Those criteria is what we call for the purpose of this thesis a "Smart" method. We precise that the word "Smart" is here for naming purpose. The construction follows the same structure as the "Random" N portfolio. Based on the entire dataset returns, we choose first the asset that fits the most the criterion and then build growing portfolios on the remaining assets based on the criterion, with the last portfolio of 100 N assets containing our original dataset. In this case we only have one simulation per criterion.

Now that we have described the functioning of our selection, let us talk about the criteria we chose. As mentioned earlier, we wanted to select criteria based on an investor choice. In our context, we wanted to select criteria based on simple decisions that an investor could use to make his choices. We selected three : first, a choice base on the historical average returns of the dataset. Second, a choice based on the historical variance of returns of the dataset. Third and last, one based on the momentum of the past 12 month of an asset :

1. The choice of the historical average returns of the dataset seemed evident to us as it may be the most simple choice an investor could make. It can easily be calculated by anyone whose willing to invest if you have information on the historical return of the asset. The investor then choose the asset with the best average return first and the worst in last.

2. Regarding our second choice, the historical variance of returns is also an easy choice that an investor could make. If we hold the second assumption of Markowitz seen in section 2.2, a rational investor sees variance of return or volatility as an undesirable thing. Therefore, he chooses in priority the asset with the lowest variance.

3. Last, we decide to select a choice that would represent another type decision making. Indeed, it can happen that investors are sometimes tempted to choice an asset not based on the entire historical average return but rather on the recent trend of the asset.If an asset has had very good returns in recent times, an investor might be tempted to invest in this one rather than one who had the same historical average return but performed less well in recent times. As our historical returns span is long, we decided to select only the average return for the last 12 months.
   Therefore, the investor in this case chooses first the asset that has the best momentum for the last 12 months and the one who has the lowest momentum in last. This choice is far from only being psychological. Indeed, Jegadeesh & Titman (2001) show that there are substantial evidences that indicates that an asset that performs best (or worst) over a three to 12-month period tend to continue to perform well(or poorly) over the next three to-12month period. In other words, this means that an asset with a high momentum tends to perform better then an asset with a low momentum.

3.2.3 Portfolio strategy and choice of weights

To evaluate "Random" N and "Smart" N both individually and together. We need to decide what portfolio strategies we will use regarding the choice of weights within one portfolio. As we wanted to have more than one strategy to evaluate, we decided to use 4 portfolios strategies that have been presented on the section 2.4, 2.5 and 2.6 of this thesis. We use them to observe the different values that optimal N may take be depending on the strategy but also if our 2 methods presented in the subsection above have an impact on the strategy. We choose the Sample mean-variance portfolio (SMV), the Minimum-variance portfolio (minV), the Equally weight portfolio (EW) and Kan & Zhou two fund rule (KZ). For the sake of information and clarification, we put in parenthesis the initials the strategies take on graphics and results to avoid repetition or confusion. That being said, The SMV build weights on assets based on equation (18), the minV on equation (21) and the EW on equation (22). Regarding Kan & Zhou's two fund rule (KZ), we use the optimal weights given by (27). However, there is one subject we need to discuss regarding the last equation. Indeed, the optimal value of the combination coefficient given by equation (31). To use this equation out-of-sample, we need to find an estimator of θ^2. The first that comes to mind is the in-sample estimator

$$\hat{\theta}^2 = \hat{\mu}^\top \hat{\Sigma}^{-1} \hat{\mu} \tag{35}$$

But the issue with this estimator is that it is very biased and leads to estimation errors. To correct that, Kan & Zhou (2007) proposed an adjusted estimator of the latter :

$$\hat{\theta}_a^2 = \frac{(T-N-2)\hat{\theta}^2 - N}{T} + \frac{2(\hat{\theta}^2)^{\frac{N}{2}}(1+\hat{\theta}^2)^{-\frac{T-2}{2}}}{TB_{\hat{\theta}^2/(1+\hat{\theta}^2)}(N/2,(T-N/2)} \tag{36}$$

where

$$B_x(a,b) = \int_0^x y^{a-1}(1-y)^{b-1}\,dy \tag{37}$$

is the incomplete beta function. The first part of this estimator is the unbiased estimator of θ^2 and the second part of the estimator is the adjustment to improve the unbiased estimator when it is too small but also to assure that $\hat{\theta}^2$ is not negative which is desirable as θ^2 is always positive. Regarding the value of T and N, we see in the details of our rolling window explanation that the value of T is equal to 120 while N is simply equal to the number of asset present in the portfolio evaluated. For the value of γ, we perform all optimizations and calculations assuming $\gamma = 1$ as this thesis will not attempt to address risk aversion selection method.

3.2.4 Rolling window and output

To assess out-of-sample performance on our different methods (Random and Smart) and portfolio strategy (SMV, minV, EW and KZ), we evaluate them according to their out-of-sample utility using a rolling window technique. The rolling window works as follows : to evaluate out-of-sample performance we need to choose a window of a certain length that represents our time of estimation. This window represent the in-sample returns we base ourselves on to evaluate our out-of-sample performance. In our case, we choose a 10-year windows (120 months). Each 120-month window generates a specific portfolio based on the strategy chosen and this same portfolio is applied to the 121st. As a first output, we have then an out-of-sample return based on the optimal weights calculated on the in-sample window (January 1970 to December 1979) multiplied by the returns of the 121th period (January 1980). This operation is repeated as the window moves in time. In the end, the rolling window gives us 507 out-of-sample returns from (January 1980 to March 2022). To calculate to out-of-sample utility, we simply use the utility function in equation (2).

For the Random N method, it means that we should repeat this operation for each portfolio (100 times) as N is growing ,for each simulation (100 times),for each portfolio strategy (4 times) and for each dataset (3 times). To realise the scale of this operation it means that we need to calculate 120000 out-of-sample utility. Through experimentation of the computational power of our equipment to realise analysis, we realised that the computation of 1 simulation took 4 minutes on average. If we scale to the total, it means that we should run 80 hours straight of computation. This outcome being extremely inefficient in terms of time. We decided to find a way to lower the number of computation while remaining relevant for our analysis. This is why we choose to effectuate jumps in N to calculate out-of-sample utility. Indeed, instead of having 100 portfolios, we choose to select portfolio for N (2,5,10,15,...) up to $N = 100$. This allows us to divide our number of computation by 4 and therefore having ”only” 20 hours of computation which is a lot easier to manipulate and plan. Our final output of this computation would then be 100 vectors of 21 out-of-sample utilities per strategy and per dataset which is still a considerable amount of information. To simplify the reading of the data and also the visual of graphics, we chose to compute an average

of these 100 vectors so that we only find ourselves with 1 vector per portfolio strategy and per dataset, meaning that we have 12 in total which is way easier to handle specially for graphics.
For the Smart N, we simply repeat the same rolling window technique except for the fact that, with our decisions imposed by criteria on our datasets, we don't need to do 100 simulations for each dataset which facilitate our computational work as we only require 12 simulations per criterion.

3.3 Result & Analysis

To ease the presentation of our results, we present first the results given by our ”Random” N method and analyse them.Then, we go through the results of our ”Smart” criteria and compare them with the random ones.

3.3.1 Random N selection

As explained in the methodology, our final output for this subsection is the average utility of the 100 simulations made for each dataset and portfolio strategy, with N going from 1 to 100. To illustrate the results, we plot the 4 graphics below which are regrouping the average utility of each dataset for a specific portfolio strategy : Figure 7 for the Sample Mean-Variance, Figure 8 for the minimum Variance, Figure 9 for the Equally weighted and Figure 10 for the Kan and Zhou's two fund rule.

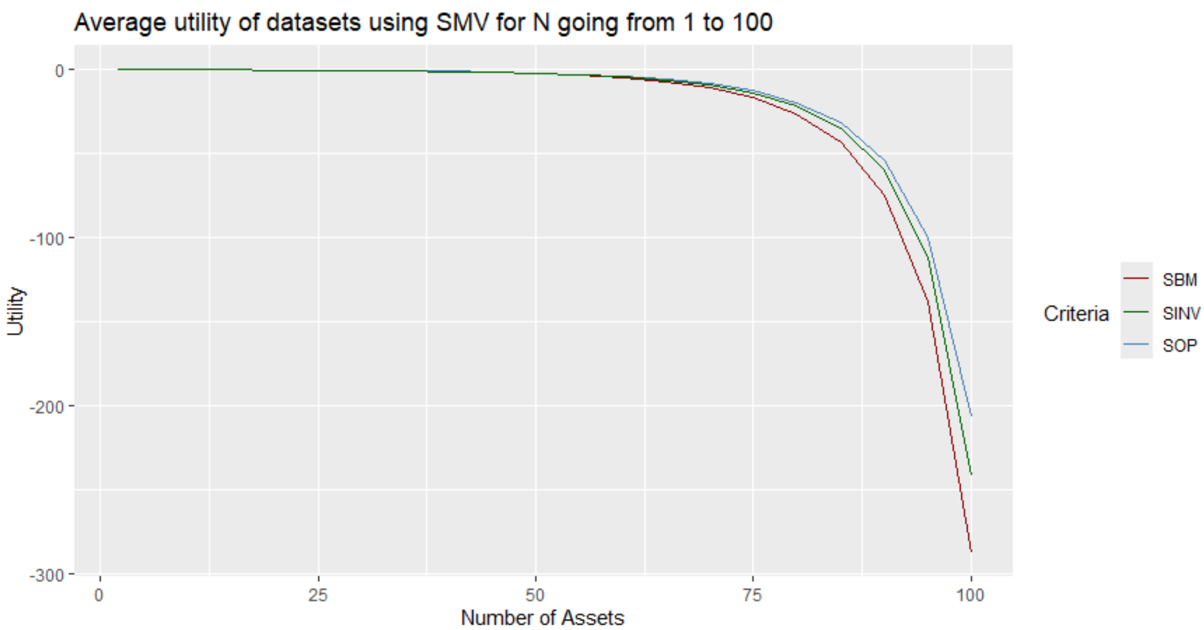

Figure 7: Average utility of the datasets using the portfolio strategy SMV for N going from 1 to 100 using the Random method

Starting with the SMV, we can see on the graphic above that we have a clear trend line regarding the optimal value of N. Indeed we see that the utility is continuously decreasing as N increase. This trend was already observed been observed on Figure 4 in our theoretical part, reassuring our statement on how exponentially poor SMV performs out-of-sample when N increase. The optimal number N assets to have in the SMV is very small. More specifically, we can see that the average utility for $N = 2$ for the datasets are respectively 0,00172 for SBM, $-4,3^{-5}$ for SOP and 0,0026 for SINV.

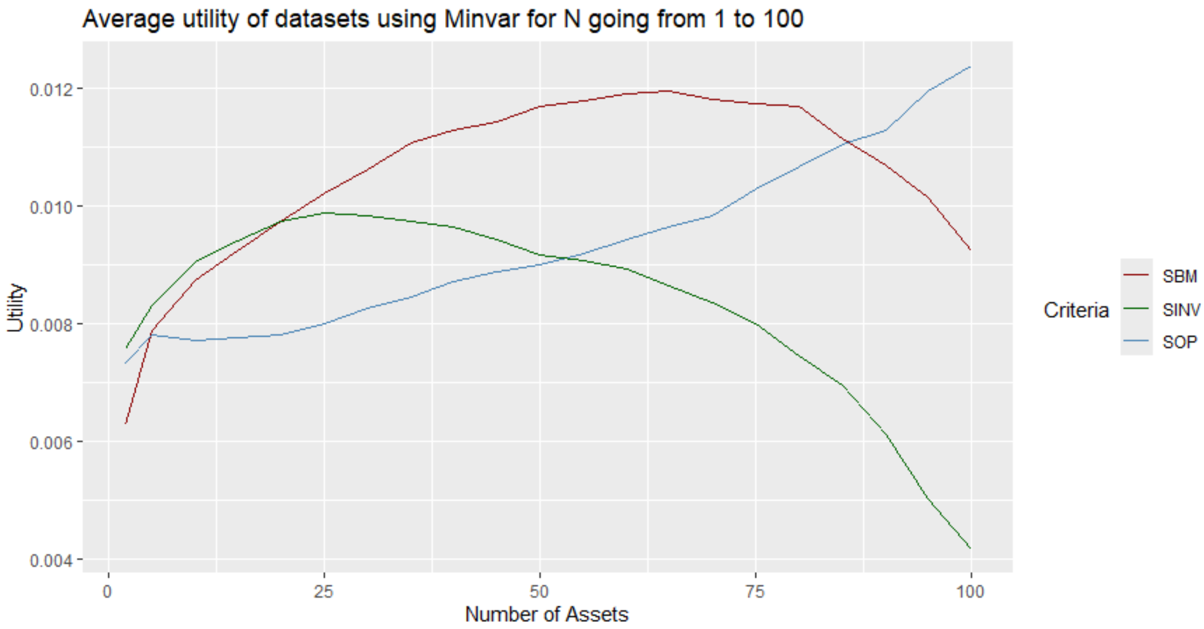

Figure 8: Average utility of the datasets using the portfolio strategy minV for N going from 1 to 100 using the Random method

Following with the minV, we can observe a very different trend line that the one we saw previously. Every dataset has his own optimal N, and have very different trend line of the evolution of N. SBM reaches is optimal value at $N = 65$ where the average utility equals 0,0112. SOP reaches is optimal value at $N = 100$ where the average utility equals 0,0124. Last, SINV reaches its optimal value at $N = 25$ where the average utility equals 0,00988. These results may mean that finding an optimal value for the minimum variance portfolio strategy is challenging has every dataset seems to have his own trend line regarding the evolution of the number N of assets in the portfolio. Why do we see these variations ? The minimum variance strategy allocates more weights to assets with low variance while completely ignoring the mean returns. A consequence of this allocation is that when an asset with a low variance but also a low mean return is added to our portfolio we could observe a decrease of the utility. On the opposite, an asset with a high mean return but also with a high mean variance can be completely ignored by the portfolio.

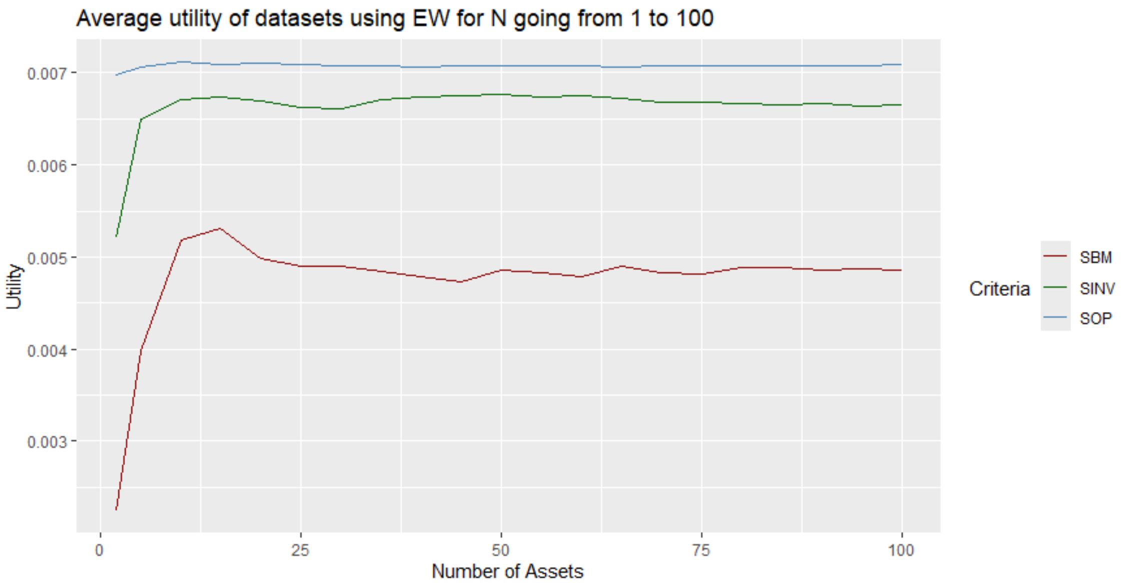

Figure 9: Average utility of the datasets using the portfolio strategy EW for N going from 1 to 100 using the Random method

Let us now move on to our naïve portfolio EW where we can observe a recurring pattern in the evolution of the number of assets. For SBM we reached to optimal value of $N = 15$ with a value of 0,0053. For SOP, we observe a flat trend line where the peak value is reached at $N = 10$ with a value of the average utility at 0,0071. Last, we have SINV having a trend line slightly different than the two others. Indeed, although he also reaches a peak at $N = 15$ of 0,0675, we have an interval of N between 45 and 60 that shares the same value. We can still remark that a pattern seems to repeat as peak values reached when the number of assets N is around 10 to 15. These results are aligned to the early findings of Evans and Archer (1968) and their optimal value of N to be around 10 or so and that returns stabilize once N is large enough.

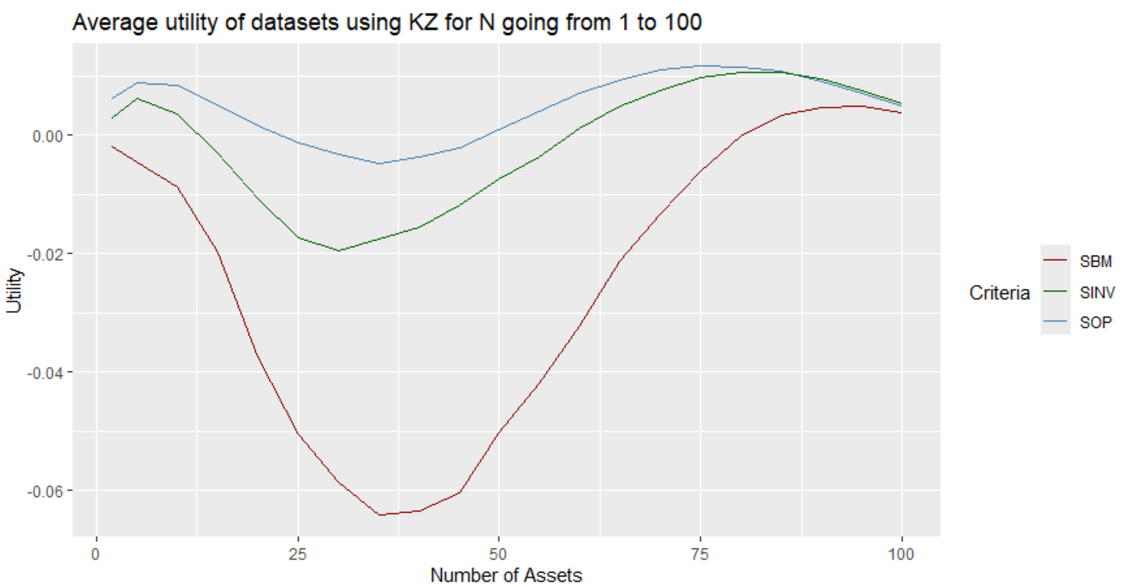

Figure 10: Average utility of the datasets using the portfolio strategy KZ for N going from 1 to 100 using the Random method

Last for our "Random" N method, we have our KZ portfolio strategy where we saw on Figure 6 that the optimal value of N should be around T/2. As T is here equal to 120 we could have expected an optimal value around 60. In our graphic above, we instead observe a trend line of a very different shape. Indeed, the average utility starts with a decreasing trend as N increase before switching trend as N reaches 35 for SBM and SOP, and 30 for SINV. We observe then an increase in utility for all dataset and reach its peak at $N = 95$ for SBM, 75 for SOP, and 80 for SINV with respective values of 0,00475, 0,0116 and 0,01. Using these results, we can suppose that the KZ portfolio would tend to be more optimal when N is between 75 and 95. This final result being different from the one of T/2 in section 2.6. It indicates still that the optimal value of N is higher than the SMV or the EW portfolio.

If we now compare the different optimal values of the average utility for each dataset, we observe that the minimum variance portfolio strategy has the best results in terms of average utility for the datasets SBM and SOP, with SINV having KZ as its best performing strategy with the minV still being very close. On the other hand, this domination of the minV is also marked by the absence of pattern in the optimal value N which in our case is not desirable as we are trying to find trend and patterns of the optimal N value for a specific portfolio strategy that can be extended to any datasets. For the rest, the KZ and EW portfolio strategies have lower average utility than the minV and stand at the second and third place. The advantage of these strategies being that there seem to be a pattern regarding the optimal value of N which is way more interesting in the context of our research. As we could imagine, the SMV is always the worst performing strategy in every dataset.

Now that we have all this information about the portfolio strategies and their respective optimal N. Let us see in next section if our 3 smart criteria influence either the value of N or the utility.

3.3.2 Smart N selection

Though we keep the same structure as the Random N section where we analysed the results per portfolio strategy, in the willingness to make our graphics clear and readable, we decided to split them per dataset.

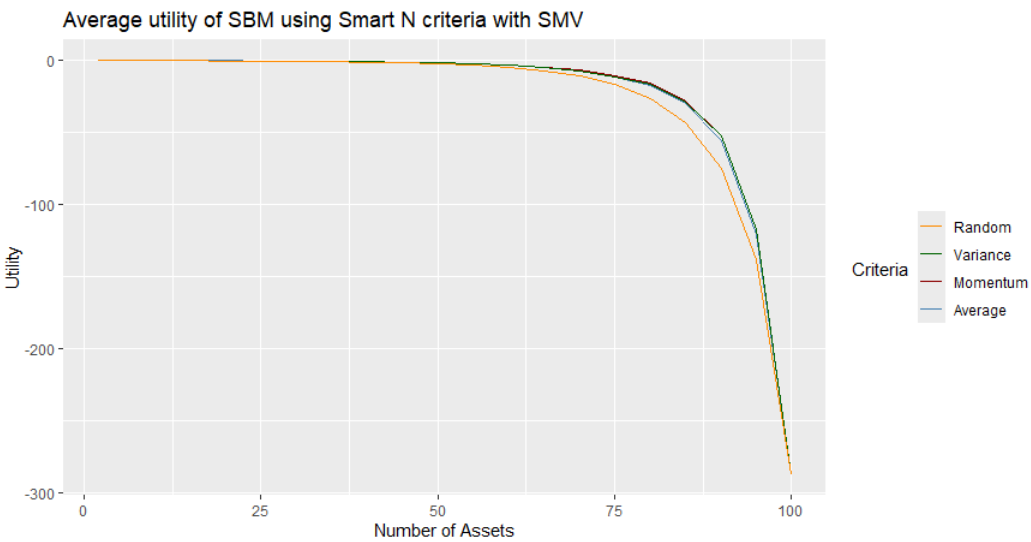

Figure 11: Average utility of SBM using Smart N criteria for N going from 1 to 100 with SMV

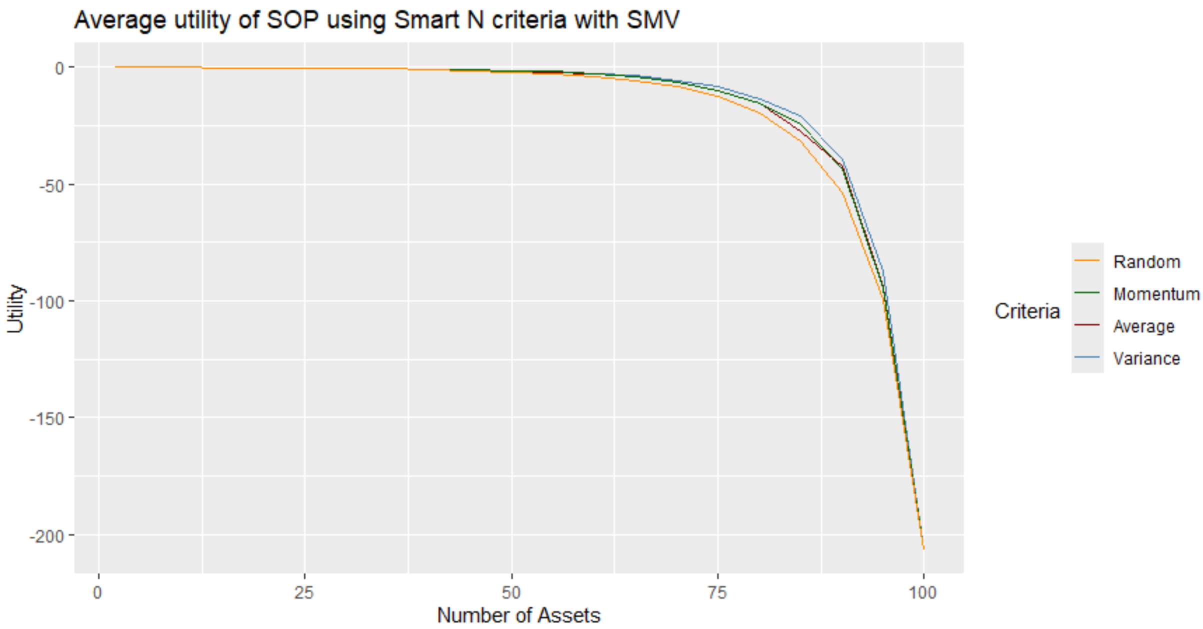

Figure 12: Average utility of SOP using Smart N criteria for N going from 1 to 100 with SMV

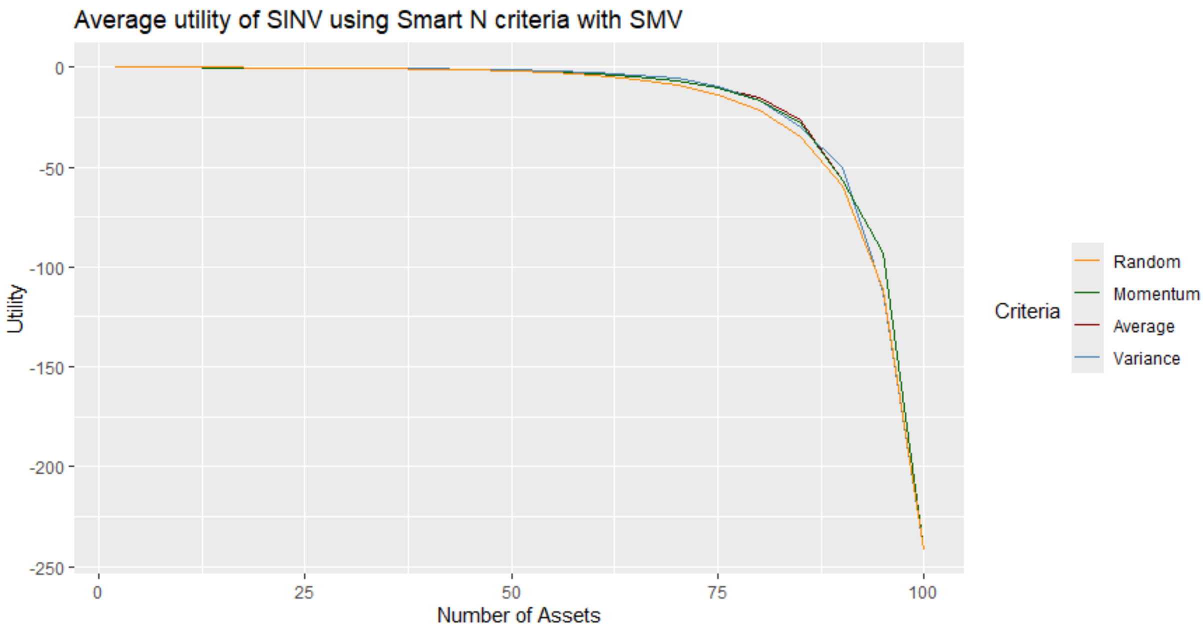

Figure 13: Average utility of SINV using Smart N criteria for N going from 1 to 100 with SMV

Coming back to our SMV portfolio, we can observe in the 3 graphs above that the smart criteria do not change the decreasing trend line of N. However, for all datasets it seems that using the criteria improves the utility of the portfolio with the variance criterion giving the best results with a utility of 0,0062 which is 6 times more than the random method.

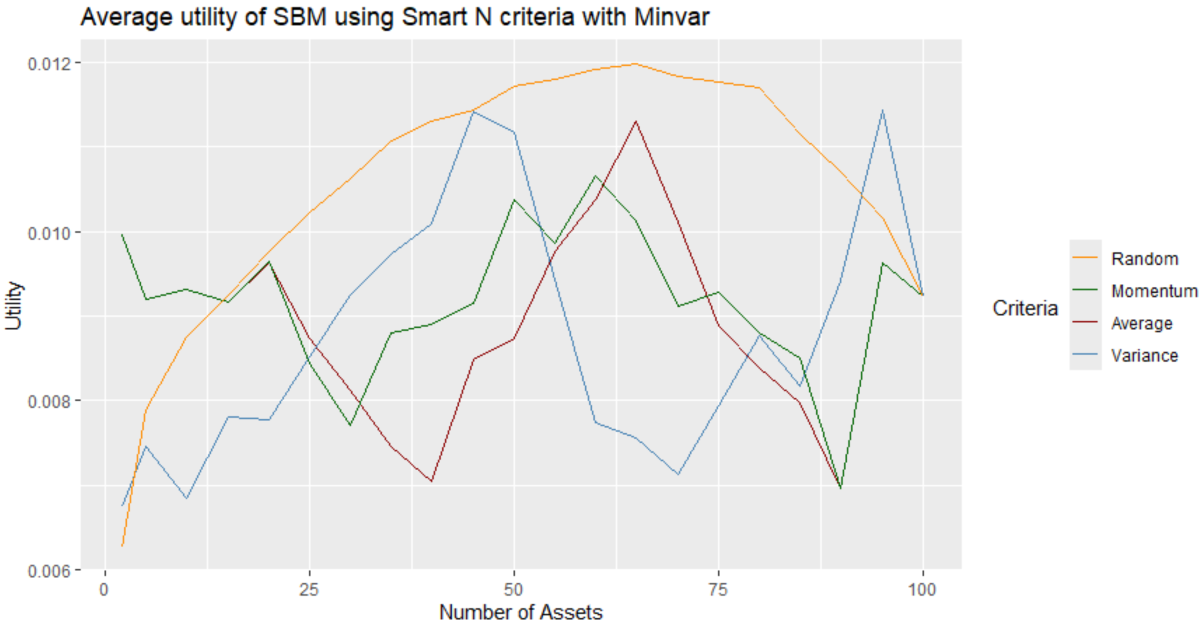

Figure 14: Average utility of SBM using Smart N criteria for N going from 1 to 100 with minV

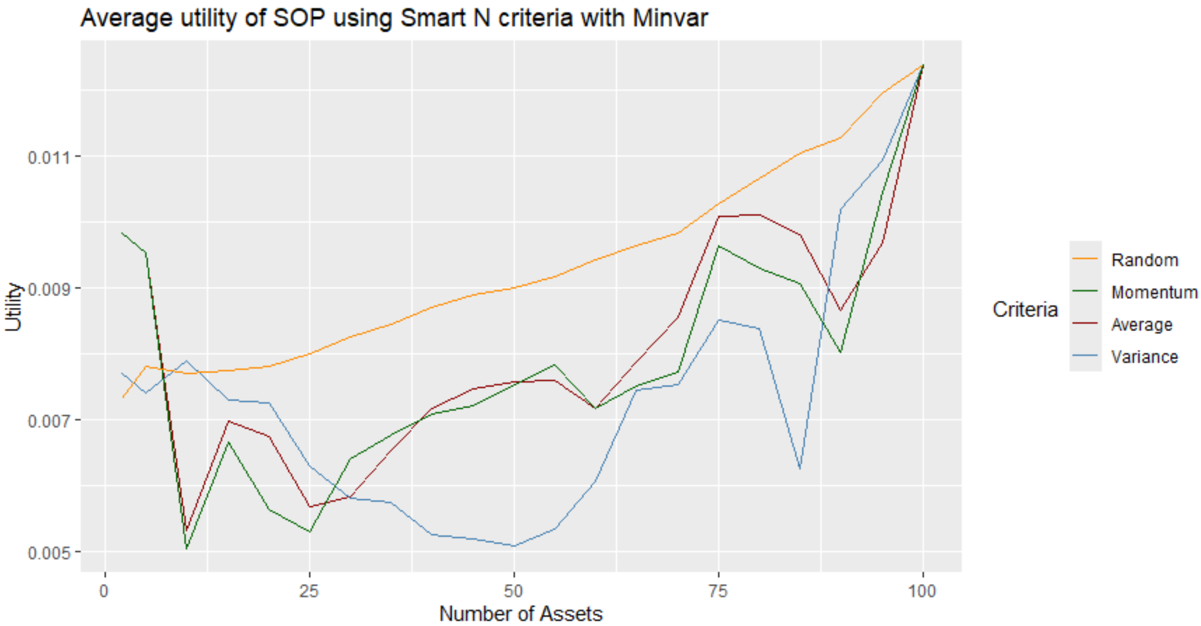

Figure 15: Average utility of SOP using Smart N criteria for N going from 1 to 100 with minV

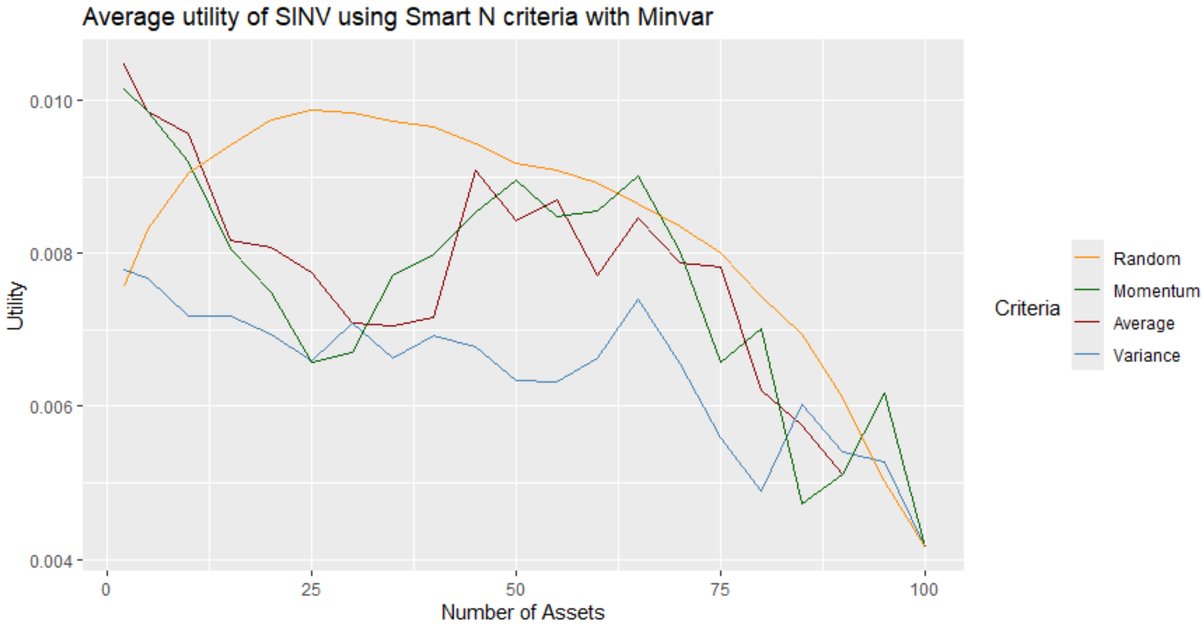

Figure 16: Average utility of SINV using Smart N criteria for N going from 1 to 100 with minV

The minimum Variance portfolio seems also to repeat its unpredictable pattern regarding the optimal N. Furthermore, even the performance of the criteria seems to variate for each dataset. For SBM, the random method performs the best at $N = 65$ while for the criteria, The average and momentum criteria seem also to reach their peak at $N = 65$ and 55 respectively while the variance criterion reaches its maximum both at $N = 45$ and 95. For SOP, the random method also stands as the best average utility reaching its peak with all criteria at $N = 100$. Last with SINV, we observe an opposite trend than SOP , with peak values being reached at $N = 2$ by criteria with average and momentum criteria performing the best. All of this confirms our previous statement regarding the irregularity and the unpredictability of the optimal N that the minimum variance portfolio had. Also, the fact that the random method is the one who performs best as N increase is unique to this strategy. This would mean that, in the context of a portfolio attributing more weights to low variance assets, selecting assets using the 3 different criteria is not optimal and that selecting assets randomly becomes the best strategy as N increases.

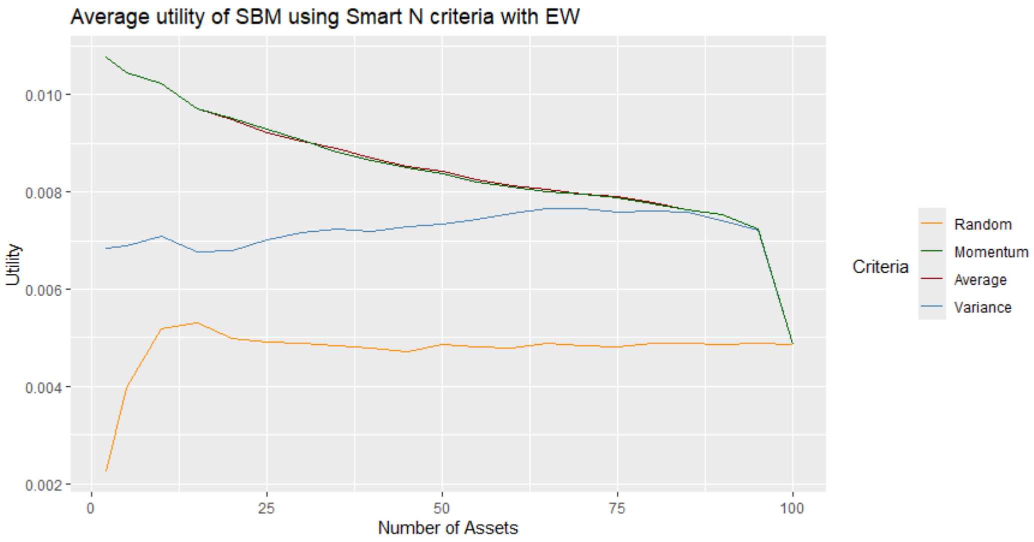

Figure 17: Average utility of SBM using Smart N criteria for N going from 1 to 100 with EW

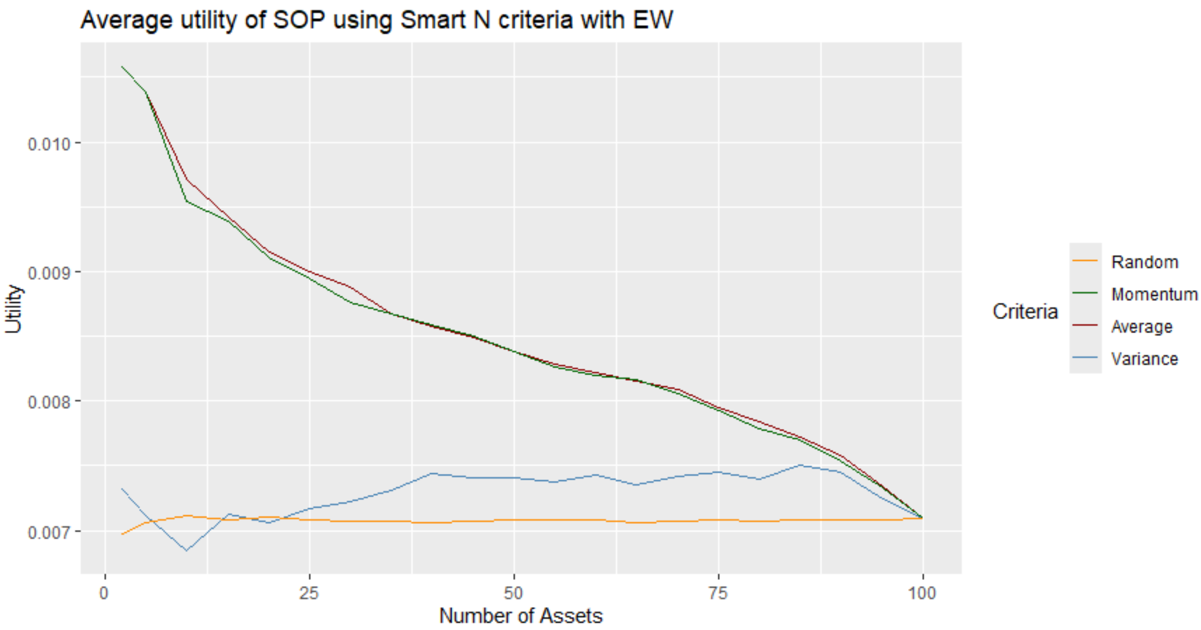

Figure 18: Average utility of SOP using Smart N criteria for N going from 1 to 100 with EW

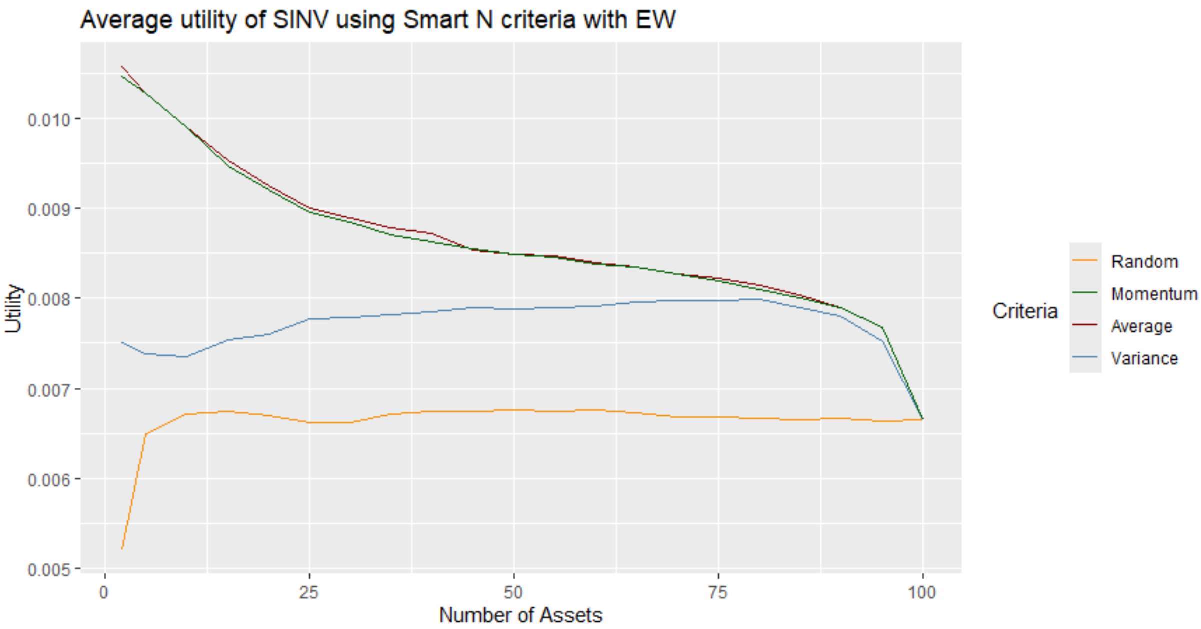

Figure 19: Average utility of SINV using Smart N criteria for N going from 1 to 100 with EW

The equally weighted portfolio combined to our smart criteria gives us interesting results. Where the random method reached its peak at $N = 15$, the average and momentum criteria here performs the best with an optimal N being small while the variance criterion, still performing better than the random method, reaches its peak when N is higher. The main lesson that we must put forward here is that the average and momentum criteria perform better than the random one and that we have a trend of N being small.

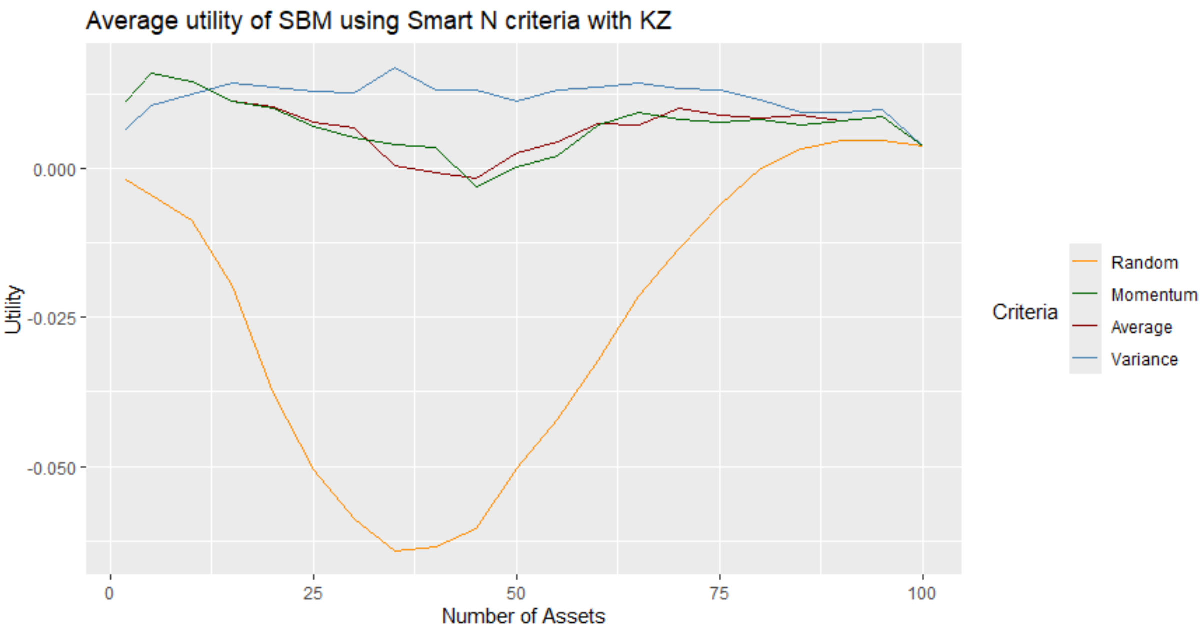

Figure 20: Average utility of SBM using Smart N criteria for N going from 1 to 100 with EW

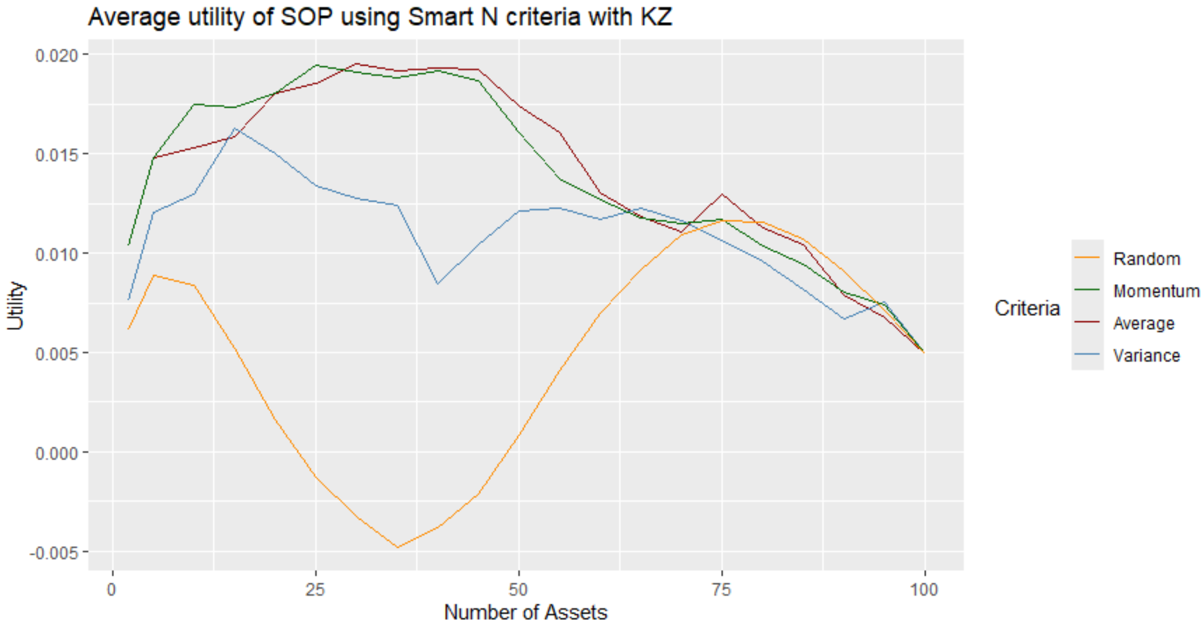

Figure 21: Average utility of SOP using Smart N criteria for N going from 1 to 100 with EW

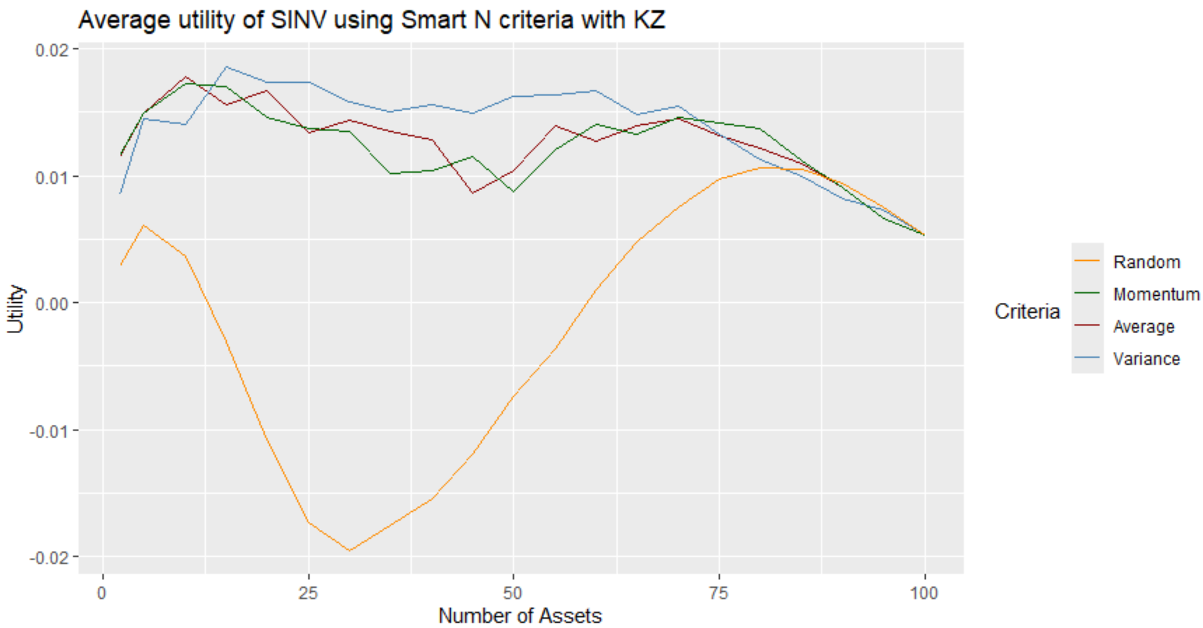

Figure 22: Average utility of SINV using Smart N criteria for N going from 1 to 100 with EW

Last with Kan and Zhou's two fund rule, the implementation of our smart criteria completely changed the trend we observed in our random part. If the trend of the optimal N was between 75 to 95, the number of assets at the best utility has considerably reduced. The first general trend of the three datasets is that the random method is again the worst performing method. Now if we look into SBM's details, we see that the best performing criterion is the variance one with a utility value of 0,0167 at $N = 35$ with the average and momentum criteria having a value of 0,0158 at $N = 5$. For SOP the average criterion performs best again at N=35 with a value of 0,0195. Last with SINV, variance is again the best performing criterion with an optimal $N = 15$ with a value of 0,0185. With these numbers on hand, it is difficult to say which criterion is the best as the average and variance came both as the best one depending on the dataset. But the fact that we have twice the value $N = 35$ and $N = 15$ suggest that if we use the criteria, it would make more sense to invest less assets than in the random method and go for a range between 15 and 35. Interestingly, now the curve for KZ is closer to that in theory, i.e., it first increases and then decreases.

Reflecting on our findings in the random method, we saw that the minimum variance portfolio performed best but was lacking of a pattern regarding the number optimal of N asset while the KZ and EW coming second and third could be observed with a certain trend for N. If we look at our smart criteria now, we see that KZ has the best utility for all datasets. On the counterpart, we saw that our trend regarding the optimal N changed drastically as we went from 75 to 95 to 15 to 35. If all the criteria performed better than the random method, we couldn't define which one was the best to use, but as KZ is always performing better than the other portfolio strategies, the choice of the criterion has not too much impact.

Coming after the KZ strategy, we have the minV and the EW standing between the second and third place. Regarding the criteria we can observe different trend. In the EW, The average and momentum criteria are clear dominant in this strategy with the random method again being the worst performing. For the minV, the irregularity seems also to spread to the criteria with the random method performing best for SBM, the average criterion for SINV and even all criteria being equal as N optimal is equal to 100. Last for the SMV, though it is the worst performing portfolio we can still mention that the variance criterion performs best and the worst being random as seen before.

Focusing on our optimal number of N assets and what we have seen in the theoretical section, we saw that the decreasing trend of the sample mean-variance portfolio was both present in theory and in our empirical research. The fact that it is the worst performing portfolio strategy is also not surprising as we already considered it as the worst performing portfolio in our theory. Regarding the minimum variance, we observed with the random method that it was the best performing portfolio but was suffering from an optimal number N of assets being unpredictable. This lead in performance disappeared when we used our smart criteria as the equally weighted was closing to the minV in terms of performance using the average and momentum criteria with the advantage that we could observe a pattern of N being equal to 1 or 2 for the EW.

Anyway, those performance were still far from the KZ portfolio. We presented the Kan and Zhou's two fund rule in detail in our section 2.6 and after computing the optimal value of the weights using a certain estimator for θ^2, we saw that the optimal value of N is theoretically equal to T/2. When we look at our empirical results, we do not have the same optimal N but at a value standing between 15 to 35 which is lower than expected but still being interesting due to the fact that we have the same curve as in theory. This difference in N may be due to the fact that we do not use the same estimator of θ^2 in the empirical section than in theory.
Nonetheless, we saw that our best performing portfolio was the KZ using the different criteria though there was not one criterion being clearly better than the others except for the momentum and random ones that were often lower than the average and variance in terms of performance.

Discussing of criteria, we saw that they always outperform the random method except for one exception for SBM using the minimum variance portfolio. How could this be explained ? This may be due of the nature of the utility function. As we subtract the the variance from the mean with a factor of 1/2 as $\gamma = 1$, a criterion that chooses the asset with the best in sample mean or the lowest in sample variance may tend to have a consequent improvement to the portfolio in the early stages as N is not high. In the SMV, as the optimal N is small, choosing the best assets first have clearly an impact on your utility.
If we saw first a trend of N being between 10 and 15 for the EW portfolio, using the criteria reduced the number of optimal assets to 2 but strongly improved the performance of the portfolio. This pattern was also present in KZ as we went from a range of the optimal N of 75 to 95 to a range of 15 to 35 but again with considerable augmentations in the performances. This could mean that using a specific criterion, in our case average, variance or momentum would lead to a reduction of the optimal N assets but improving considerably the performance of portfolio.
In our case, the best choice being to invest in the KZ portfolio with a number of asset between 15 and 35 using a criteria between the average or the variance, a choice that could be defined by the investor preference or specific to the dataset.

4 Conclusion

In the literature review, we had the opportunity to use the work of Zaimovic & al. (2021) to investigate on the theory of optimal number of N assets in one's portfolio. We discussed the pioneering work of Evans and Archer that suggested investing in maximum of 10 N assets. But going through all the other papers in the literature, we saw that the optimal N was very dependant to the strategy that was chosen to evaluate the performance of the portfolio, but also the strategy used to compute the optimal weights and even the dataset chosen to evaluate the performance. Two extreme examples were the work of Haensly (2020) and Arvanitis & al. (2023) where if the core of the research is to find an optimal N, the methodologies and means to find their results are completely different.

In the desire to focus our theoretical research, we then presented the work of Markowitz (1952) and the direct impact that his theory and assumptions have on our optimal N before introducing his famous mean-variance portfolio, where we detailed the equations to optimise the weight of the assets within the portfolio to optimise it considering that the true parameters of the model are known. We then saw that in reality it is rarely the case (Kan & Zhou, 2007). After computing the equations of the model taking the reality pf the parameter uncertainty into account, we saw that the sample mean-variance portfolio performed poorly out-of-sample and lead to an N being very small.

To solve the issue, we went through different ways to alleviate this parameter uncertainty using different portfolio strategies and parameters such as the minimum variance portfolio and the equally weighted portfolio but also and most importantly more complex methods such as the Kan & Zhou's two-fund rule which combines the sample mean-variance portfolio and the risk free asset using a specific coefficient. We developed the equations of this portfolio strategy to find that the most optimal N to maximise the utility of this portfolio was between $[0,T/2]$, T being our number of historical observations.

To put these theoretical findings, we decided to see whether they were applicable using real datasets with historical returns of assets. To diversify our research, we used two different methods to approach the problem. A "Random" N method where we used random simulations of our datasets to build portfolios of size 1 to 100, and the " Smart" N method where we used 3 different criteria to build our portfolio of size 1 to 100. These criteria dictated how an investor would choose to add an asset to his portfolio. The criteria are : first, a choice based on the best in sample average return of an asset. Second, a choice based on the lowest variance and third a choice based on the momentum of an asset. The more an asset fit the criterion the sooner he is chosen to build our growing portfolio with the worst asset being the last.

Regarding the choice of the optimal weights we tested these two methods on 4 portfolios. The sample mean-variance, the minimum variance, the equally weighted and the Kan and Zhou's two fund rule portfolios. We also developed in details how we choose to estimate the parameters Kan and Zhou's two fund rule portfolio.

To test out-of-sample performance, we evaluated the average out-of-sample utility using a rolling window technique detailed in section 3.2.4 with the size of the window being equal to 120 months. We then compared the different portfolios strategy performance for the both the random and smart method and see whether we observe patterns of an optimal number N of assets in these portfolios. In section 3.3 we went into the details of our results, where we observed that the best performing portfolio for the random method is the minimum variance but with the issue that the optimal number of asset is irregular and unpredictable. On the other side, we have the equally weighted and Kan and Zhou portfolios, that perform less than the minimum variance but have pattern regarding the optimal value of N being 10 to 15 and 75-95 respectively. When comparing our smart criteria to our random results. We observed that the criteria always improves the performance of every portfolio strategy except for the minimum variance. Furthermore, we saw that the Kan and Zhou portfolio becomes the best performing portfolio for every dataset. We also had a decrease in the number of optimal assets becoming 2 for the equally weighted and 15 to 35 for Kan and Zhou. For every tests, the sample mean-variance is always the worst performing portfolio.

With all this information in mind, we can now reflect on our work and discuss it, especially the empirical approach. We discussed many factors that could have influenced our final statement of the optimal N, with the first being our dataset. As explained in the dedicated section, we only worked with 3 different datasets, this work can be extended to different datasets to see whether we can observe the same trends or not.
More in details of our methodology, different parameters such as the T value of the rolling window, gamma, could have been tried with different values to see if it may have an impact or not. Same with the portfolio strategies used, other strategies presented or mentioned in the theoretical part could have been used to see whether the Kan and Zhou portfolio would remain the best choice or not. Another point that could be discussed is how the smart criteria are chosen. Indeed, when selecting which assets fits the best, we use the history of the entire database. But, this method introduces a look-ahead bias. An investor, who starts to invest at the 121st period of the rolling window, does not have access to the future information that we have. Ideally, one could repeat this thesis's test with the smart criteria only based on the first 120th returns of the dataset. Furthermore, one could also introduce a turnover of the asset has the rolling window advance as the best asset for a criterion could change in time.

Ultimately the reason of theses limitations are coming from a constraint that we mentioned in the thesis : the computational power. Indeed, we already mentioned in our methodology that the numbers of simulation created by the random method was the main reason we limited our work. At first, this thesis also included a fourth dataset being the historical returns of the S&P 500. But as we wanted to expand our methods to this dataset we realised that our computational power was clearly limited as soon as the number of N assets within a dataset or portfolio was above 100, having an exponential computational time increasing as N was increasing. This is also the reason why we limited our number of simulations to 100. Maybe that an increase of the number of simulations could improve or worsen the performance of the random method compared to the smart criteria. About the criteria, though we chose to do only 3 but our code and methods could be used with different criteria depending on the investor's need. Then, one could expand this thesis's work by either increasing the total number of simulations or datasets, either by testing new criteria.

Finally I think that the question of the optimal N number of assets is a vast domain that is very interesting and filled with unanswered questions that remain to be solved. Nonetheless, if we had to answer to our question "How many assets should one select to optimize out-of-sample performance?", we could state that, based on the results of our empirical research crossed with our theoretical review, an investor should invest in 15 to 35 assets using the Kan & Zhou's two fund rule portfolio strategy and should select his assets based on a criterion of average or variance.

5 Bibliography

References

[1] Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10.2307/2975974

[2] Arvanitis, S., Scaillet, O., & Topaloglou, N. (2023). *Sparse spanning portfolios and under-diversification with second-order stochastic dominance.*.Document non publié, Athens University of Economics and Business, Athènes.

[3] Zaimovic, A., Omanovic, A., & Arnaut-Berilo, A. (2021). How Many Stocks Are Sufficient for Equity Portfolio Diversification? A Review of the Literature. *Journal of Risk and Financial Management*, 14:551. doi:10.3390/jrfm14110551

[4] Evans, J. L. & Archer, S. H. (1968). Diversification and the Reduction of Dispersion: An Empirical Analysis. *The Journal of Finance*, *23*(5), 761–767. https://doi.org/10.2307/2325905

[5] Haensly, P. (2020). Risk decomposition, estimation error, and naïve diversification. *The North American Journal of Economics and Finance* 52: 101146.https://doi.org/10.1016/j.najef.2020.101146

[6] Kan, R. & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, *42*(3):621–656. doi:10.1017/S0022109000004129

[7] Ledoit, O. & Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, *88*(2):365-411. https://doi.org/10.1016/S0047-259X(03)00096-4

[8] Hull (2018).*Risk Management and Financial Institutions*, 5th Edition— Instructor Companion Site. (s.d.).https://bcs.wiley.com/hebcs/Books?action=index&itemId=1119448115&bcsId=11147

[9] Sharpe, W. F. (1966). Mutual Fund Performance. *The Journal of Business, 39*(1), 119–138. http://www.jstor.org/stable/2351741

[10] Black, F. (1993). Estimating Expected Return. *Financial Analysts Journal*, *49*(5), 36-38. https://doi.org/10.2469/faj.v49.n5.36

[11] Best, M.J. & Grauer, R.R. (1991) On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results. *The Review of Financial Studies*, *4*, 315-342. https://doi.org/10.1093/rfs/4.2.315

[12] DeMiguel, V., & Nogales, F. F. (2007). Portfolio Selection With Robust Estimation. *Social Science Research Network*. https://doi.org/10.2139/ssrn.911596

[13] Michaud, R. O. (1989). The Markowitz optimization enigma: Is “optimized” optimal? *Financial Analysts Journal*, *45*(1), 31–42.http://dx.doi.org/10.2139/ssrn.2387669

[14] Lassance, N. (2022).*Big data in Finance : Robust linear Regression.*Document non publié, Université catholique de Louvain, Louvain-la-Neuve.

[15] DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, *22*(5):1915–1953. doi:10.1093/rfs/hhm075

[16] Jagannathan, R., & T. Ma. (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *Journal of Finance*, *58*:1651-84. https://doi.org/10.1111/1540-6261.00580

[17] Stein, C. (1956). Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution. in *3rd Berkeley Symposium on Probability & Statistics 1*, 197-206.

[18] Stein, C. (1975). Estimation of a covariance matrix. In *Riesz Lecture, 39th Annual meeting.*Atlanta, Georgia : Institute of Mathematical Statistics

[19] Tu, J., & Zhou, G. (2011). Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics, 99*(1), 204–215. https://-doi.org/10.1016/j.jfineco.2010.08.013

[20] Ledoit, O., & Wolf, M. (2004b). Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management, 30*(4), 110–119. https://doi.org/10.3905/jpm.2004.110

[21] Lassance, N., Martín-Utrera, A., & Simaan, M. (2024). The Risk of Expected Utility Under Parameter Uncertainty. *Management Science, 0*(0). https://doi.org/10.1287/mnsc.2023.00178

[22] Kan, R., & Lassance, Nathan. (2024). Optimal Portfolio Choice with Fat Tails and Parameter Uncertainty. Available at SSRN. http://dx.doi.org/10.2139/ssrn.4652814

[23] Kan, R., Wang, X., & Zhou, G. (2021). Optimal Portfolio Choice with Estimation Risk: No Risk-Free Asset Case. *Management Science, 68*(3):2047-2068. https://-doi.org/10.1287/mnsc.2021.3989

[24] Statman, M. (1987). How Many Stocks Make a Diversified Portfolio? *The Journal of Financial and Quantitative Analysis, 22*(3), 353–363. https://doi.org/10.2307/2330969

[25] Elton, E. J., & Gruber, M. J. (1977). Risk Reduction and Portfolio Size: An Analytical Solution. *The Journal of Business, 50*(4), 415–437. http://www.jstor.org/stable/2352173

[26] Statman, M. (2004). The Diversification Puzzle. *Financial Analysts Journal, 60*(4), 44-53. doi:10.2469/faj.v60.n4.2636

[27] Lee, Y., Kim, W.C., Kim J.H. (2020). Achieving Portfolio Diversification for Individuals with Low Financial Sustainability. *Sustainability, 12*(17):7073. https://doi.org/10.3390/su12177073

[28] Jegadeesh, N., & Titman, S. (2001). *Momentum.*Document non publié, University of Illinois and Business, Illinois. http://dx.doi.org/10.2139/ssrn.299107

[29] Kenneth R. French website for the datasets:
https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html