Optimal Covariance Matrix with Parameter Uncertainty for the Tangency Portfolio

Ghislain Yanou[†]

June 3, 2009

Abstract

We provide a model for understanding the impact of the sample size neglect when an investor, hoping for the tangency portfolio uses the Sharpe model and the sample estimator of the covariance matrix for this purpose. By assuming the null hypothesis, we compute the expected loss of the sample covariance matrix. The null hypothesis being actually wrong regarding the market structure of returns, we consider the opposite of the expected loss, and assume that it is the expected loss between the sample covariance matrix with the true covariance matrix. Although this value is not really the true one, we use it in order to catch the dynamic between the true covariance matrix and the sample one through time. We compute the risk aversion of investors having some specific ambiguity reliance on the Sharpe model and obtain a set of covariance matrices characterizing the ambiguity and the risk aversion of investors. We highlight a knock-on phenomenon of the covariance matrix and provide a covariance matrix for a risk management issue and a covariance matrix for an allocation decision under uncertainty by deriving an equilibrium state between investors having a maximum ambiguity and investors having a minimum ambiguity reliance on the Sharpe model. We show that ambiguity comes actually from the sample size of the investment universe and follows a power law distribution. A preliminary study in the S&P500 universe seems confirm our results.

Keywords: Ambiguity, covariance matrix, null hypothesis, power law, representativeness, tangency portfolio.

J.E.L. Classification: G.11, G.12, C.11.

[*]We are grateful to Thierry Chauveau, Emmanuel Jurczenko for help and encouragement in preparing this work. We would also like to acknowledge the Office of International Program of Brown University (USA) for their financial supports. The content of this article engages only its authors and does not necessarily reflect the opinions of their employers.

[†]University of Paris-1 (Finance/CES/CNRS), 106 bd de l'hôpital F-75647 Paris cedex 13 - Office 419. Phone: (+33)1-44-07-42-71, Fax: (+33)1-44-07-82-47. E-mail: ghislain.yanou@malix.univ-paris1.fr.

Optimal Covariance Matrix with Parameter Uncertainty for the Tangency Portfolio

Abstract

We provide a model for understanding the impact of the sample size neglect when an investor, hoping for the tangency portfolio uses the Sharpe model and the sample estimator of the covariance matrix for this purpose. By assuming the null hypothesis, we compute the expected loss of the sample covariance matrix. The null hypothesis being actually wrong regarding the market structure of returns, we consider the opposite of the expected loss, and assume that it is the expected loss between the sample covariance matrix with the true covariance matrix. Although this value is not really the true one, we use it in order to catch the dynamic between the true covariance matrix and the sample one through time. We compute the risk aversion of investors having some specific ambiguity reliance on the Sharpe model and obtain a set of covariance matrices characterizing the ambiguity and the risk aversion of investors. We highlight a knock-on phenomenon of the covariance matrix and provide a covariance matrix for a risk management issue and a covariance matrix for an allocation decision under uncertainty by deriving an equilibrium state between investors having a maximum ambiguity and investors having a minimum ambiguity reliance on the Sharpe model. We show that ambiguity comes actually from the sample size of the investment universe and follows a power law distribution. A preliminary study in the S&P500 universe seems confirm our results.

Keywords: Ambiguity, covariance matrix, null hypothesis, power law, representativeness, tangency portfolio.

J.E.L. Classification: G.11, G.12, C.11.

Optimal Covariance Matrix with Parameter Uncertainty for the Tangency Portfolio

1 Introduction

An investor making an optimal decision needs to know the true parameters under the asset returns. But the true parameters are unknown by the decison maker and have to be estimated. Under the framework of Markowitz (1952)[53], investor cares only about expected returns and volatility of static portfolio and should hold the tangency portfolio on the efficient frontier. Following Tobin (1958)[72] and Sharpe (1964)[70], the optimal portfolio for investor is the mix between the tangency portfolio and the risk-free asset (see Brandt (2004)[6] for a survey of the academic literature about the mean-variance framework). To implement this portfolio in practice needs to estimate both expected returns and covariance matrix from the time series. The standard estimators used by practitionners and academic research for this purpose are the sample mean for the expected returns and the sample covariance matrix for the covariance matrix.

However, Mandelbrot (1963)[52] pointed out the insufficiency of the normal distribution for modeling assets returns. As a result, the sample tangency portfolio obtained from the Sharpe (1964)[70] model, due to estimation errors in the sample mean and sample covariance matrix, produces extreme weights that fluctuate substantially over time and performs poorly *out-of-sample* (see Michaud (1989)[56], Best and Grauer, (1991)[4] and Litterman (2003)[49]). But as Meucci (2005)[55] stated, the sample tangency portfolio remains the most important benchmark model because the sample parameters are as well nonparametric estimators as maximum likelihood estimators.

Two approaches[1] have been developed by literature for dealing with the poor performances of the the sample tangency portfolio: the *plug-in* method and the decision theory. While the *plug-in* method consists to specify parameters of the datas generating process and plug them into the Markowitz framework, the aim of the decision theory is to build a portfolio allocation process coherent with the subjective view or preferences of investor.

[1]We may cite alternative approaches as the resampling method of Michaud (1998)[57] and Michaud and Michaud (2008)[58], but this approach is so computational that makes impossible to explain the behaviour of investor. Concerning robust portfolio allocation rules, see Goldfarb and Iyengar (2003)[26] and Garlappi *et al.* (2007)[24]. For portfolio with moment restrictions, see for instance MacKinlay and Pástor (2000)[51].

It is well known that the expected returns is more difficult to estimate than the covariance matrix (see for instance Merton (1980)[54]), and that errors into the sample mean have a larger impact in the expected *out-of-sample* performances than errors into the sample covariance matrix, see Chopra and Ziemba (1993)[9]. Following Brandt (2004)[6], it seems that the uncertainty about the estimation of the expected returns tends to increase the posterior variance of the distribution, and the uncertainty about the estimation of the variance tends to fatten the tails of the posterior distribution. Therefore, a well estimator of the covariance matrix is helpful for taking into account the deviation of asset returns from the normal distribution.

From the *plug-in* view, several approaches have been proposed to deal with the problem of estimating the elements into the covariance matrix. One approach consists to impose some good properties in the sample estimator of the covariance matrix by shrinking the empirical covariance matrix. Ledoit and Wolf (2004-a[47]) propose a weighted average estimator of the covariance matrix between the sample covariance and the identity matrix having a better condition number. Fan *et al.* (2007)[22] use a similar approach for giving a stationary property to a time-domain estimator of the covariance matrix.

A second approach consists to give some structural properties to the covariance matrix by imposing a portfolio norm constraint (see Frost and Savarino (1988)[23] and Chopra (1993)[10]). DeMiguel *et al.* (2007)[13] suggest to impose a norm constraint on the portfolio allocation program and show some analytical relations between this constraint and the shortage threshold which can be supported by investor. Following Jagannathan and Ma (2003)[35], it seems this approach rewards on shrinking the sample covariance matrix.

Several empirical evidences call into question the one factor model (see Fama and French (2004)[21]) and show that except the market factor, others risk factors exist and should be taken into account (see Black *et al.* (1972)[5]), this is at the origin of the multi-factor models. Some statistical methods like the principal component analysis have been used by the literature to extract factors in the historical returns, but this approach does not allow for distinguishing factors which contain real information from the noise. The Random matrix theory developed by physicians in order to understand the energy process for which sources are unknown (see Edelman (1989)[17]), gives a solution for filtering noise. Laloux *et al.* (1999)[45] show some empirical evidences justifying the use of the Random matrix theory in finance. Following the Edelman's thesis (1989)[17], Plerou *et al.* (2002)[63] perform a study of the Random matrix theory to understand cross-correlation of the high frequency financial returns. A recent work on the Random matrix theory applied in finance comes from Potters *et al.* (2005)[64], Conlon *et al.* (2008)[11] and Yanou (2008)[74].

An alternative method to understand moments of a distribution is obtained by a linear combination of order statistics named L-moments (see Hosking (1990)[33] and Serfling and Xiao (2007)[68]). Introduced by Sillito (1951)[71] and popularized by Hosking *et al.* (1985)[29], L-moments can be interpreted, like classical moments, as simple descriptors of the shape of a general distribution and they offer a number of advantages over conventional moments (see Hosking (1986,[30], 1989[32] and 1990[33]) and Hosking and Wallis (1987)[31]). Serfling and Xiao (2007)[68] develop co-Lmoment in a multivariate framework. Yanou (2008)[74] proposes an estimator of the covariance matrix for the global minimum variance portfolio, based on the L-moments and shows that the Random matrix theory can be used to extract information from noise.

Another approach to estimate the covariance matrix consists to use robust statistics as M-estimators or S-estimators (see for instance DeMiguel and Nogales (2007)[15]), or the minimum determinant covariance matrix (see Rousseeuw and Van Driessen (1999)[66]).

The estimation methods above, only focus on the estimation of the covariance matrix without take into account the subjective view or preference of investor. The decision theory is another literature dealing with the poor properties of the sample tangency portfolio. Brandt (2004)[6] list three ways for this purpose. We only list two ways because of the similarity of one among them with the *plug-in* approach.

The first way consists to eliminate dependence of the optimization process with the true parameters by replacing them with a subjective view or a prior distribution of investor. This is a Bayesian approach[2] widely used by literature for dealing with the estimation errors into the sample tangency portfolio. Following Brown (1978)[7], the sample tangency portfolio, or more specifically the two-fund rule which is the mix between the sample tangency and the risk-free asset, is generally outperfomed by the Bayesian decision rule under a diffuse prior. Kan and Zhou (2007)[43] give a theoretical demonstration of this result and show that a large class of two-fund rules are outperformed by the Bayesian rule. They propose a three-fund rule, which is a mix between the two-fund rule and the global minimum variance portfolio. It seems that this approach outperfomed the Bayesian rule. However the three-fund rule of Kan and Zhou (2007)[43] have some limits. Actually authors do not deal with the case where there is a shortsale constraint in the optimization process. Following Frost and Savarino (1988)[23] and Jagannathan and Ma (2003)[35], we know that the introduction of a shortsale constraint in the portfolio optimization program improves the performances of the sample covariance matrix, a fortiori the performances of

[2]For a purely statistical approach, see for instance Barry (1974)[2], Bawa *et al.* (1979)[3]. For an approach based on shrinkage estimators, see for instance Jobson *et al.* (1979)[37], Jobson and Korkie (1980)[36], and Jorion (1985[38], 1986[39]) for the estimation of expected returns and Ledoit and Wolf (2003[46], 2004-b[48]) for the estimation of the covariance matrix. For an approach that rely on an asset-pricing model for establishing a prior, see for instance Pástor (2000)[61] and Pástor and Stambaugh (2000)[62].

the sample tangency portfolio. In addition, Kan and Zhou (2007)[43] remains in an ideal framework where returns follow an *i.i.d.* normal distribution.

The second way comes from the behavioral finance literature (see Barberis and Thaler (2003)[1] for a survey on the behavioral finance) where the aim consists to explain financial puzzles or inefficiency in the markets by assuming that some agents are not fully rational. Following some psychological beliefs[3] of investor, the Von-Neumann and Morgenstern (1944)[73] approach based on the maximization of the expectation of a utility function has been rejected. Several non-expected utility models have been proposed for explaining the behavior of investor under uncertainty when the probability relative to the outcomes of investor choices are known. We may cite among them, the disappointement theory (see Gul (1991)[27], Chauveau and Nalpas (2009)[8]), the regret theory (see Loomes and Sugden (1982)[50]) and the prospect theory of Kahneman and Tversky (1979)[41]. The prospect theory is popular in the financial literature because of its ability for explaining some financial puzzles, but in the case of decision making, Savage (1964)[67] proposes a subjective expected utility approach where investor weighted the utility function by his (her) subjective probability. But the experiment of Ellsberg (1961)[18] calls into question this model by introducing an ambiguity aversion of people into the gamble. For instance, an investor hoping for the tangency portfolio and using the sample estimators for this purpose, may take into account the non-gaussian nature of asset returns and the finite sample size character of the investment universe, and then the nature of information. Following Kahneman and Tversky (1974)[40], it seems that people often use the representativeness heuristic well explained in Cont (1998)[12] in the case of estimation of parameters in finance. One of the biais of the representativeness heuristic is the sample size neglect or the law of small number (see Rabin (2002)[65], where investor thinks that small sample will reflect the properties of the parent distribution.

The aim of this article is to propose a model for explaining the impact of the sample size neglect when investor hopes for the tangency portfolio and uses in this context the sample covariance matrix. For this purpose we build a set of covariance matrices characterizing the ambiguity that investor should have. We then derive some specific covariance matrices and show the existence of an equilibrium among them. This equilibrium is not obtained by the criterion widely used by literature and based on the maximization of the expected utility. Our approach seems consistent with the Robust control literature (see Gilboa and Schmeidler (1989)[25], Hansen and Sargent (2001)[28]).

The remainder of this paper is organized as follows. Section two presents the general framework. In section three, after explaining the theoretical process, we provide an expression for the covariance matrix and show some properties of this one. In section four, we

[3] For details about these beliefs, see Barberis and Thaler (2003)[1].

provide some specific covariance matrices and highlight the representativeness heuristic. We provide an equilibrium state in section five and show that the ambiguity comes from the size of the investment universe. In section six, we show the consistency of the model with the market structure by using assets from the S&P500 universe and apply the model to the portfolio optimization. We conclude the paper in section seven.

2 The General Framework

We start by defining some notations and hypothesis. Consider the standard portfolio choice problem of an investor who cares only about the expected returns and the volatility of his (her) portfolio from an universe of N risky assets. Let x the vector of portfolio weights invested in the N risky assets available in the investment universe. We assume that μ and $\mathbf{\Omega}$ denote respectively the true $N \times 1$ vector of expected returns and the true $N \times N$ covariance matrix from the universe $\mathbf{R}$ of size $T \times N$ matrix of excess returns[4]. We assume that the elements of the vector of expected returns are different of zero[5], and the optimal portfolio weight x^* invested in the risky assets is obtained by the maximization of the following utility function:

$$U(x) = x^{\mathbf{T}}\mu - \frac{\gamma}{2}x^{\mathbf{T}}\mathbf{\Omega}x \tag{1}$$

where γ denotes the risk aversion of investor with $\gamma > 0$. The solution of the maximization of the utility function is given as follows:

$$x^* = \frac{1}{\gamma}\mathbf{\Omega}^{-1}\mu \tag{2}$$

Actually, the covariance matrix is unknown and need to be estimated. We then have an utility function depending only of an estimator of the covariance matrix:

$$U(\hat{\mathbf{\Omega}}) = \frac{1}{2\gamma}\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}^{-1}\mu \tag{3}$$

Following the ability for an estimator to well estimates the covariance matrix, investor increases or decreases his (her) utility and expected utility functions. The aim for investor is to use an estimator of the covariance matrix maximizing his (her) utility function. Let $\hat{\mathbf{\Omega}}_s$ denotes the sample covariance matrix. We introduce the expected loss function of the sample covariance matrix $\hat{\mathbf{\Omega}}_s$, where $\mathbf{\Omega}$ is the true covariance matrix:

$$\rho\left(\mathbf{\Omega}, \hat{\mathbf{\Omega}}_s\right) = U(\mathbf{\Omega}) - E\left[U(\hat{\mathbf{\Omega}}_s)\right] \tag{4}$$

[4]The excess returns are obtained by removing to the asset returns the risk-free rate.

[5]By this assumption, we avoid a gaussian distribution for returns.

The true covariance matrix being unknown, we can not compute the expected loss function of the sample covariance matrix. But, if we assume that the true covariance matrix is known it then becomes straightforward to compute this value. We propose an approach to resolve this issue consisting to consider a wrong covariance matrix. By inverting the dynamic between the sample covariance matrix and this wrong covariance matrix, we may understand the dynamic between the sample covariance matrix and the true covariance matrix. In the next section, we provide a solution to this issue.

3 The Set of Covariance Matrices for the Tangency Portfolio

The true covariance matrix is unknown, that makes impossible to compute the expected loss function of the sample covariance matrix. Remember, our aim is to explain the impact of the small sample size neglect of investor. Assume we use the identity matrix as the true covariance matrix. The structure of the identity matrix is well explained in Meucci (2005)[55]. While the sample covariance matrix denotes a symmetrical ellipse centered on the expected returns, the identity matrix may be viewed as a circle centered on the expected returns. The non-Gaussian nature of the returns in markets may be characterized as an non-symmetrical ellipse having some outliers. Although the identity matrix knows some good statical properties as its well conditioning, from a geometrical point of view, the sample covariance matrix fits better the shape of the returns in the market. As a result, by assuming the null hypothesis, we consider that the identity matrix is the true covariance matrix, which is actually wrong regarding the ellipse characterizing asset returns, however it becomes straightforward to compute the expected loss function of the sample covariance matrix.

The null hypothesis is the basis of the application of the Random matrix theory in finance. Yanou (2008)[74] proposes an estimator of the covariance matrix obtained from the multivariate L-moments (Serfling and Xiao (2007)[68]) and filtered using the Random matrix theory. Author shows that the resulting portfolio minimizing the volatility[6] well performs the sample global minimum variance portfolio when a shortsale constraint is imposed[7].

However, DeMiguel *et al.* (2007)[14] show that the sample based mean-variance strategy and its extensions need around 3000 months for a portfolio with 25 assets and 6000 months for a portfolio of 50 assets to outperform the naive strategy. Their results let

[6] Characterized by the multivariate L-moment of order two.

[7] The results of Jagannathan and Ma (2003)[35] shows that the sample covariance matrix becomes robust when a shortsale constraint is imposed.

think that the use of the identity matrix is simply optimal or at least more optimal than strategies developed by literature. Actually in their empirical study, authors use simulated datas from the one factor model, datas corresponding to several sectorial indices and datas coming from the Fama and French (1993)[20] portfolios.

In the one factor model, noise into the naive allocation is simply the deviation of the market weights of the underlying assets from the average. The beta of each asset characterizing their risks and their cross-correlations are not at all taken into account. The identity matrix does not contains cross-correlation errors, as a result may be a well estimator of the covariance matrix than several estimators in this case. Furthermore, it is well known that sectorial indices are less correlated than stocks, and by considering only sectorial indices, authors design datas for which the identity matrix may be the true covariance matrix[8]. We can use the same argument when they consider many portfolios from Fama and French (1993)[20], which corresponds to different factors which are by design statistically independents. The following section focus on the theoretical process.

3.1 The Theoretical Process

By assuming the null hypothesis, it becomes straightforward to compute the asymptotical expected loss function of the sample covariance matrix, but keep in mind that is a wrong hypothesis. The expected loss function of the sample covariance matrix is then defined as follows:

$$\rho\left(\mathbf{I}, \hat{\mathbf{\Omega}}_s\right) = U(\mathbf{I}) - E\left[U(\hat{\mathbf{\Omega}}_s)\right] \tag{5}$$

From the expression of the utility function of the investor, we have:

$$U(\mathbf{I}) = \frac{1}{2\gamma}\mu^{\mathbf{T}}\mu \tag{6}$$

and:

$$E\left[U(\hat{\mathbf{\Omega}}_s)\right] = \frac{1}{2\gamma}\mu^{\mathbf{T}}E(\hat{\mathbf{\Omega}}_s^{-1})\mu \tag{7}$$

The following proposition gives the expected behavior of the sample covariance matrix in this case:

Proposition 1. If the identity matrix is the covariance matrix, the expected loss function of the sample covariance matrix ρ is defined as follows:

$$\rho = -\frac{(\mu^{\mathbf{T}}\mu)(N+2)}{2\gamma(T-N-2)} \tag{8}$$

Proof. See appendix 141.

[8]The constant-correlation model of Elton and Gruber (1973)[19] may be used for building a well estimator of the covariance matrix in this case.

Under the null hypothesis, the identity matrix is the covariance matrix and then, ρ is the expected loss function of the sample covariance matrix. Remember, the null hypothesis is actually wrong regarding the market structure. If we assume that the identity matrix has a higher expected loss function with the covariance matrix than the sample one, we may understand the behavior between the true covariance matrix and the sample one by inverting the asymptotical expected loss function obtained with the wrong hypothesis. Therefore $-1/\rho$ can be viewed as the expected loss function between the true covariance matrix and the sample one:

$$\rho\left(\mathbf{\Omega}, \hat{\mathbf{\Omega}}_s\right) = \varrho \tag{9}$$

where ϱ is equal to $-1/\rho$:

$$\varrho = \frac{2\gamma(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{10}$$

The parameter ϱ can be viewed as the starting theoretical expected loss function between the true covariance matrix and the sample covariance matrix. Remember, we try to understand the dynamic between the true covariance matrix and the sample one. The parameter ϱ is not actually the true value of the expected loss of the sample covariance matrix, but the starting point for understanding its dynamic with the true covariance matrix. Since we just trying to understand the dynamic of this parameter we do not need to have its true value. For instance, at the present time, we assume under the null hypothesis that the expected loss of the sample covariance matrix is equal to -2. At the next time, it is equal to -3. Therefore, the sample covariance matrix is closer of the identity matrix at the next time. As we assume that the null hypothesis is a wrong hypothesis, we consider that the expectd loss of the sample covariance matrix with the true one is equal to $1/2$ at the initial time and $1/3$ at the next time, that is, the sample covariance matrix is closer of the true one, when it is farther of the identity matrix. We do not focus on the true value of the expected loss, but on its dynamic.

We then may compute a set of covariance matrices depending on the risk aversion parameter by setting the equality between the expected loss function $\rho\left(\mathbf{\Omega}, \hat{\mathbf{\Omega}}_s\right)$ and ϱ. In the next sub-section, we give an expression of the covariance matrix.

3.2 An Expression of the Covariance Matrix: The General Case

By assuming that ϱ is the starting expected loss function of the sample covariance matrix it becomes straightforward to compute an expression of the covariance matrix satisfying the starting value of the expected loss of the sample covariance matrix. The following proposition gives an expression of the covariance matrix obtained:

Proposition 2. A candidate $\tilde{\mathbf{\Omega}}$ for the covariance matrix is defined as follows:

$$\tilde{\mathbf{\Omega}} = \hat{\mathbf{\Omega}}_{\mathbf{s}} + \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{4\gamma^2(T-N-2)}\mathbf{\Lambda} \tag{11}$$

where $\mathbf{\Lambda}$ the matrix of size $N \times N$ obtained from the vector of expected returns μ:

$$\mathbf{\Lambda} = \mu \otimes \mu^{\mathbf{T}}$$

with the sign $\otimes$ denoting the kronecker product.

Proof. See appendix 242.

It is straightforward to see that the matrix of expected returns $\mathbf{\Lambda}$ is a symmetric matrix. As a result, $\tilde{\mathbf{\Omega}}$ is also a symmetric matrix. The following proposition shows that $\tilde{\mathbf{\Omega}}$ is also a positive defined matrix:

Proposition 3. The covariance matrix obtained as a sum of the sample covariance matrix and a weighted matrix of expected returns is a defined positive matrix.

Proof. See appendix 344.

Notice that, the covariance matrix obtained is not a linear combination between the sample one and a target matrix as for the shrinkage covariance matrices of Ledoit and Wolf (2003[46], 2004-a[47] and 2004-b[48]). According to the value of the parameter α, the covariance matrix may increases or reduces its elements toward the sample covariance matrix. Actually when a linear combination between the sample covariance matrix and a target is used, this is referent to the Bayesian approach, where investor has a believe characterized by the target covariance matrix, and makes an arbitrage between the estimation errors into the sample covariance matrix which is asymptoticaly unbiaised and the biais into the target matrix which contains less estimation errors. Our approach allows of understanding the dynamic between the sample covariance matrix and the true covariance matrix for an investor hoping for the the tangency portfolio.

The covariance matrix obtained adds to the sample one, the matrix $\mathbf{\Lambda}$ weighted by a parameter α defined as follows:

$$\alpha = \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{4\gamma^2(T-N-2)} \tag{12}$$

The parameter α depends on two parameters; the risk aversion parameter γ and the size of the investment universe. As a result, according to the risk profile of investor, there is an unique covariance matrix and there is an infinite number of covariance matrices corresponding to the infinite number of risk aversion profile of investors.

Since the risk aversion parameter γ is positive, if the number of historical returns T is smaller than $N-2$, the parameter α is negative. In this case, the sample covariance matrix adds up the covariance matrix with the matrix $\mathbf{\Lambda}$. That is, the elements of the sample covariance matrix need to be shrunk. This observation is consistent with the shrinkage approach of Ledoit and Wolf (2003)[46]. As authors state, when the number of stocks N is of the same order of magnitude as the number of historical returns T, the total number of parameters to estimate is of the same order as the total size of the data set, which is clearly problematic. It implies that the most extreme coefficients in the matrix thus estimated tend to take on extreme values not because this is the truth, but because they contain an extreme amount of error, furthermore it is necessary to reduce these coefficients. Michaud (1989)[56] calls this phenomenon *error-maximization*, and when N goes to infinite, the sample covariance matrix tends to have more extreme elements.

When $(N+2) < T$, the parameter α is positive. In this context, the sample covariance matrix actually reduces the elements of the covariance matrix. Notice that, higher the number of historical returns and smaller the magnitude of reduction of the elements and when T goes to infinite, the sample covariance matrix is the same with the covariance matrix. This observation is consistent with the asymptotical behaviour of the maximum likelihood estimators which are unbiaised asymptotically.

Assume now that the number of historical returns T increases of ΔT, therefore α turns down of $d_T(\alpha)$ with:

$$d_T(\alpha) = \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{4(\gamma+\Delta\gamma)^2(T-N-2)^2}\Delta T \tag{13}$$

In this case, the sample covariance matrix tends to move closer from the covariance matrix. When the risk aversion parameter γ decreases of $\Delta\gamma$, from time T to time $T+\Delta T$, the parameter α increases of $d_\gamma(\alpha)$ where:

$$d_\gamma(\alpha) = \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{2\gamma^3(T+\Delta T-N-2)}\Delta\gamma \tag{14}$$

In this case, the sample covariance matrix tends to move away from the covariance matrix, and will be the same according to the gap between $d_T(\alpha)$ and $d_\gamma(\alpha)$. If investor decreases his (her) risk aversion from time T to time $T+1$, the gain of the sample covariance matrix due to the increase on the sample size may be canceled by the change in the risk aversion profile of investor. Therefore, in order to remains the gap between the covariance matrix and the sample one constant, $d_T(\alpha)$ must be at least higher than $d_\gamma(\alpha)$. This observation may be helpful for understanding the behavior of the sample covariance matrix as T increases. Following the popular believe, the central limit theorem, as T goes to the infinite, the sample covariance matrix must be the true one. However, the increase of the number of historical returns does not necessarily improves the efficiency of the sample covariance matrix. The following proposition states this observation:

Proposition 4. As the number of historical returns increases from T to $k\Delta T$, the sample covariance matrix goes toward the true covariance matrix from the point $T+k\Delta T$

if the following condition is true:

$$\langle \Delta\gamma_k \rangle \leq \frac{k\Delta T}{h}$$

with:

$$h = \frac{2(\gamma + \Delta\gamma_k)^2 (T - N - 2)^2}{\gamma^3 (T + k\Delta T - N - 2)} \tag{15}$$

where $\langle \cdot \rangle$ denotes the absolute value, $\Delta\gamma_k$ the variation of the risk aversion between $T + (k-1)\,\Delta T$ and $T + (k-1)\,\Delta T$,.and k an integer denoting the number of period.

Proof. See appendix 445.

Following the previous proposition, $T + k\Delta T$ is the time from which the sample covariance matrix begins to go toward the true covariance one as the number of historical returns increases through k. As long as the variation of the risk aversion of investor $\Delta\gamma_k$ in absolute value is higher than $k\Delta T/h$, the sample covariance matrix will tend to go away from the true covariance matrix, in this case the central limit theorem is no more true. When k goes to the infinite, the parameter h goes toward zero, as a result $k\Delta T/h$ tends to go toward infinite and we may expect that the variation of the risk aversion parameter $\Delta\gamma_k$ is most likely smaller. However, we may have some case where the variation of $\Delta\gamma_k$ is higher than $k\Delta T/h$.

The parameter h can be viewed as the required holding period or the feasible period for which the allocation computed today remains relevant. Investor should not buy (sell) the resulting allocation beyond (before) this time. In the case where investor wants to keep the feasible period constant (for instance, an investor hoping for a holding period of one month), he (she) should change his (her) risk profile with respect to the change of ΔT. On another hand, if investor hopes for a constant risk aversion ($\Delta\gamma = 0$), he (she) should adjust the size of the feasible period h with respect to ΔT. Finally, if investor hopes for a constant ΔT, the change on the risk aversion parameter allows for adjusting the change on the feasible period.

If we assume that from time T to time $T+1$, investor becomes more risk averse, then $\Delta\gamma$ increases and as a result, investor tends to reduce the upper bound of the holding period for the portfolio. On another hand, when investor is less risk averse, $\Delta\gamma$ decreases and the upper bound of the holding period increases. In this case, higher the parameter h, and higher the difference between $d_\gamma(\alpha)$ and $d_T(\alpha)$. The sample covariance matrix moves away from the covariance matrix. However the increase of the number of historical returns T influences more the decrease on α from $d_T(\alpha)$ than the increase from $d_\gamma(\alpha)$, and as a result, tends to cancel the gap between them and when T goes to infinite, the sample covariance matrix is the same than the covariance matrix.

Kondor *et al.* (2002)[44] point out this observation by computing the noise into several measures of volatility. They show that, when the number of historical returns is large in comparison to the number of assets in the investment universe, the level of noise into the standard deviation is the same than the noise into severals robust estimators of the volatility. Actually Proposition 4 characterizes the behaviour of the investor with respect to the risk aversion. Higher the risk aversion parameter, lower the holding period of the optimal portfolio, and lower the risk aversion parameter, higher the holding period of the optimal portfolio.

Proposition 4 may also be interpreted as follows: when the risk aversion of investor is constant, its variation is null. Therefore the previous proposition may be interpreted as the interval of time for which investor do not have any aversion, whatever happens in the market. This time interval is therefore denoted by $\Delta T/h$.

In the next sub-section, we focus on the bahavior of the utility of investor with respect to the risk aversion parameter and the size of the investment universe.

3.2.1 The Utility Function of the Covariance Matrix

The utility obtained from the covariance matrix is the sum between the utility obtained from the sample covariance matrix and a parameter taking into account the size of the investment universe and the risk aversion profile of investor:

$$U(\tilde{\mathbf{\Omega}}) = U(\hat{\mathbf{\Omega}}_s) + \frac{2\gamma(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{16}$$

Since the second component of the expression above is positive, the utility from the covariance matrix is higher than the utility from the sample covariance matrix for the same level of risk aversion. This is true when the number of historical returns in the universe is higher than the number of assets ($N+2<T$), and when T goes to infinite, the following proposition gives the bahavior of the utility function:

Proposition 5. When the number of historical returns T goes to infinite, whe have:

$$U(\tilde{\mathbf{\Omega}}) = U(\hat{\mathbf{\Omega}}_s) \rightarrow +\infty \tag{17}$$

Proof. See appendix 546.

As a result, for the same level of risk aversion, the utility from the covariance matrix is higher than the utility obtained from the identity matrix. It is straightforward to see this result by reexpressing $U(\tilde{\mathbf{\Omega}})$ as:

$$U(\tilde{\mathbf{\Omega}}) = U(\hat{\mathbf{\Omega}}_s) + \frac{1}{U(\mathbf{I})}\frac{T-N-2}{N+2} \tag{18}$$

Since $U(\mathbf{I})$ is positive[9] and $(N+2)<T$, we have:

$$U(\hat{\mathbf{\Omega}}_s) < U(\tilde{\mathbf{\Omega}})$$

[9]Which is true because of the positivity of the risk aversion parameter.

and the upper bound for $U(\mathbf{I})$ is defined as follows:

$$U(\mathbf{I}) < \frac{T-N-2}{N+2} < U(\tilde{\boldsymbol{\Omega}}) \tag{19}$$

On another hand, when $T < (N+2)$, the utility from the covariance matrix is lower than the utility from the sample covariance matrix. The utility from the identity matrix is always positive. In order to keep positive $U(\tilde{\boldsymbol{\Omega}})$ the following condition should be true:

$$-\frac{T-N-2}{N+2} < U(\mathbf{I})U(\hat{\boldsymbol{\Omega}}_s) \tag{20}$$

However, the utility function characterizes the preference of investor which may turn out wrong in an *out-of-sample* framework. The expected utility characterizes what investor should gain in future with respect to his (her) risk profile today. This is the aim of the next sub-section.

3.2.2 The Expected Utility Function of the Covariance Matrix

The expected utility function from the identity matrix is always the same and is defined as follows:

$$E\left[U(\mathbf{I})\right] = \frac{\mu^{\mathbf{T}}\mu}{2\gamma} \tag{21}$$

The covariance matrix is obtained by inverting the expected loss function of the sample covariance matrix when a the null hypothesis is assumed. Since $\tilde{\boldsymbol{\Omega}}$ is the covariance marix, the expected utility of the sample covariance matrix is defined as follows:

$$E\left[U(\hat{\boldsymbol{\Omega}}_s)\right] = \frac{T}{T-N-2}U(\tilde{\boldsymbol{\Omega}}) \tag{22}$$

and the expected utility from the covariance matrix is defined as follows:

$$E\left[U(\tilde{\boldsymbol{\Omega}})\right] = E\left[U(\hat{\boldsymbol{\Omega}}_s)\right] + \frac{2\gamma(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{23}$$

By replacing the expected utility expression for $\hat{\boldsymbol{\Omega}}_s$, we then obtain:

$$E\left[U(\tilde{\boldsymbol{\Omega}})\right] = \frac{T}{T-N-2}U(\tilde{\boldsymbol{\Omega}}) + \frac{2\gamma(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{24}$$

We have another expression for the expected utility of the covariance matrix by replacing $U(\tilde{\boldsymbol{\Omega}})$ by its expression:

$$E\left[U(\tilde{\boldsymbol{\Omega}})\right] = \frac{T}{T-N-2}U(\hat{\boldsymbol{\Omega}}_s) + \frac{2\gamma(2T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{25}$$

Since $T < (N+2)$, from (23) we know that the expected utility from the sample covariance matrix is higher than the expected utility from the covariance matrix for the same level of risk aversion. If (20) is false, the utility obtained from the covariance matrix is negative, and as a result, from (22), the expected utility from the sample covariance matrix is positive. When the number of assets in the investment universe goes to infinite, the expected utility from the covariance matrix goes toward $-2\gamma(\mu^{\mathbf{T}}\mu)^{-1}$. The expected utility from the sample covariance matrix in this case decreases toward zero. The identity matrix then, gives to the investor the highest expected utility. Consequently, when the number of historical returns is not enough in comparison with the number of assets in the investment universe, the data returns used for estimating the covariance matrix are not enough relevant. But if $(N+2) < T$ the data returns become relevant. The following proposition states this result:

Proposition 6. When $T < (N+2)$, the data returns are not enough relevant for the estimation of the covariance matrix, as a result, the naive allocation is the best way for investor. When $(N+2) < T$, the covariance matrix gives to the investor the highest expected utility than the sample covariance matrix for the same level of risk aversion.

Proof. See appendix 646.

When the number of historical returns goes to infinite, the expected utility from the covariance matrix goes toward infinite, and the expected utility from the sample covariance matrix also goes toward $U(\tilde{\mathbf{\Omega}})$ because of their equality to the infinite. We see above that when $T < (N+2)$, the naive allocation is better than the one obtained from the covariance matrix. We give an explanation to this result, in the next sub-section.

3.2.3 The Knock-on Phenomenon of the Covariance Matrix

The elements into the sample covariance matrix $\hat{\mathbf{\Omega}}_s$ are the variances of assets on the diagonal and the covariances between assets out of the diagonal. Let X_1 and X_2 two random variables, the following expressions denote the variance of X_1 and the covariance between X_1 and X_2:

$$\begin{cases} var\left(X_1\right) = E\left(X_1^2\right) - E\left(X_1\right)^2 \\ cov(X_1, X_2) = E\left(X_1 X_2\right) - E\left(X_1\right) E\left(X_2\right) \end{cases} \tag{26}$$

The variance of X_1 and the covariance between X_1 and X_2 are constituted each one of two components. The first component is an order two moment which is $E\left(X_1^2\right)$ for the variance of X_1 and $E\left(X_1 X_2\right)$ for the covariance between X_1 and X_2. The second component is a squared order of the first moment which is $E\left(X_1\right)^2$ for the variance of X_1 and $E\left(X_1\right) E\left(X_2\right)$ for the covariance between X_1 and X_2. That is, the expressions for the variance of X_1 and the covariance between X_1 and X_2 also depend on the expected returns components respectively through $E(X_1)^2$ and $E(X_1)E(X_2)$. As a result, there is some errors into the elements of the sample covariance matrix coming from the errors into the expected returns. This result is consistent with the observation of Brandt (2004)[6].

Notice that elements into the matrix $\mathbf{\Lambda}$ correspond to the expected returns components of the elements into the sample covariance matrix. As a result, when the number of historical returns T is smaller than $N-2$, the parameter α is negative and the covariance matrix for the tangency portfolio tends to increase the expected returns components of the elements into the sample covariance matrix, because of the extreme coefficients into the sample covariance matrix[10]. But in the same time, as Merton (1980)[54] shows, it is more difficult to estimate expected returns than variance, as a result by trying to reduce the extreme coefficients into the sample covariance matrix, the covariance matrix actually improves the impact of errors in the estimation of the expected returns. This is a knock-on phenomenon; by trying to reduce its extreme coefficients, the covariance matrix improves the effects of the errors coming from its expected returns components. This phenomenon explains several limits of the sample covariance matrix advanced by literature. For instance, Michaud (1989)[56], Pafka and Kondor (2004)[60] state that the sample covariance matrix is ill-conditioned[11], Potters et *al.* (2005)[64] state that the eigenvalues of the sample covariance are for a large part random (unstable), Ledoit and Wolf (2003)[46] speak about the curse of dimensionality of the sample covariance matrix. These limits comes actually from this knock-on effect.

When N goes to infinite, the variance and the covariance elements of the covariance matrix are defined as follows:

$$\begin{cases} var_{\tilde{\mathbf{\Omega}}}(X_1) = var_{\hat{\mathbf{\Omega}}_s}(X_1) - E(X_1)^2 = E(X_1^2) - 2E(X_1)^2 \\ cov_{\tilde{\mathbf{\Omega}}}(X_1, X_2) = cov_{\hat{\mathbf{\Omega}}_\mathbf{s}}(X_1, X_2) - E(X_1)E(X_2) = E(X_1X_2) - 2E(X_1)E(X_2) \end{cases} \tag{27}$$

where $var_{\tilde{\mathbf{\Omega}}}(.)$, $cov_{\tilde{\mathbf{\Omega}}}(.,.)$, and $var_{\hat{\mathbf{\Omega}}_s}(.)$, $cov_{\hat{\mathbf{\Omega}}_\mathbf{s}}(.,.)$ denote respectively the variance and covariance coefficients into the covariance matrix $\tilde{\mathbf{\Omega}}$, and the variance and covariance coefficients into the sample covariance matrix $\hat{\mathbf{\Omega}}_s$.

We now know that the covariance matrix depend on the risk aversion parameter, and the size of the investment universe, we propose in the next section to explain how the representativeness heuristic characterizes by the sample size neglect impact the allocation of investors.

[10]As Ledoit and Wolf (2003)[46] state, when the number of assets N is of the same order of magnitude as the number of historical returns T, the elements into the sample covariance matrix tends to take extreme values and contain a lot of errors.

[11]A little change in the elements into the sample covariance matrix lead to a big change in the optimal portfolio.

4 Representativeness Heuristic and Ambiguity

The sample size neglect is one of the biais of the representativeness heuristic. Since, the covariance matrix depends on the risk aversion of investor and the size of the investment universe, the covariance matrix takes into account these two facts. In this section, we propose to highlight the relation between the sample size neglect (therefore the size of the investment universe), the risk aversion of investor and his (her) ambiguity reliance on the sample tangency portfolio. Following the paper of Kahn and Sarin (1988)[42], we know some implications about the ambiguity of people under uncertainty. First, decision makers consider ambiguity when making choice under uncertainty. Secondly, the attitude of people toward ambiguity varies. Next, people are willing to pay for difference in ambiguity. Finally the mean and variance alone may not account completely for choices. Concerning the latter implication, in order to consider the higher moments in their models, authors take into account the entire distribution of the second order probabilities.

We know that, the covariance matrix obtained from our model is a sum between the sample covariance matrix and a weighting of the matrix of expected returns $\mathbf{\Lambda}$. In the last section, we see that the elements into the matrix $\mathbf{\Lambda}$ are actually the second components of the elements into the sample covariance matrix. That is, our covariance matrix takes into account the entire distribution of the second moment through the parameter α. As a result, our model is consistent with the last implication of Kahn and Sarin (1988)[42]: the covariance matrix obtained from our model actually takes into account the higher moments of the distribution. The elements out of the diagonal of $\mathbf{\Lambda}$, are the products of the expected retuns of assets. As a result, when the product is negative, there is an asymmetry between the two assets and the depth of the product characterizes the extreme values into asset returns.

The first, the second and the third implications of Kahn and Sarin (1988)[42] about ambiguity, depend on the risk aversion of investor, we propose to point out these implications from our model by computing some parameters of risk aversion depending on the bahavior of investor. When facing to the Sharpe(1964)[70] model and using the sample covariance matrix for this purpose, we have from investors two extreme behaviors (the disappointment and the elation). The disapointement behavior characterizes an investor having the maximum of ambiguity for the Sharpe model when using the sample covariance matrix, and the elation behavior characterizes an investor having the minimum of ambiguity in this context. We may also have some investors with a relative ambiguity. In the next sub-section, we propose to compute the risk aversion characterizing these behaviors.

4.1 The Covariance Matrix for Investor with the Maximum Ambiguity

The covariance matrix changes with the level of the risk aversion parameter. Higher is the risk aversion parameter and more the investor is risk averse. By minimizing the utility function with respect to the risk aversion parameter, we then obtain the covariance matrix characterizing an investor hoping for the tangency portfolio and using in this context the sample covariance matrix. Our aim is then to find the risk aversion parameter for which the utility function is worst. The following expression denotes the first derivative of $U(\tilde{\mathbf{\Omega}})$ with respect to γ:

$$D_\gamma \left[U(\tilde{\mathbf{\Omega}})\right] = -\frac{\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu}{2\gamma^2} + \frac{2(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{28}$$

The utility function being a decreasing function with respect to the risk aversion parameter, the first derivative of $U(\tilde{\mathbf{\Omega}})$ with respect to the risk aversion parameter is then negative. When $(N+2) < T$, the first derivative is negative if the following condition is true:

$$-\frac{\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu}{2\gamma^2} \leq \frac{-2(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{29}$$

which means:

$$\frac{2\gamma(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \leq \frac{\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu}{2\gamma} \tag{30}$$

and by adding $U(\hat{\mathbf{\Omega}}_s)$[12] on the both sides, we obtain the following result:

Proposition 7. The utility obtained from the covariance matrix has an upper bound defined as follows:

$$U(\tilde{\mathbf{\Omega}}_g) \leq 2U(\hat{\mathbf{\Omega}}_s) \tag{31}$$

Proof. Comes from (30).

The previous proposition also shows the shape of the efficient frontier obtained from our model in comparison with the classical efficient frontier of Markowitz (1952)[53]. Therefore, a means for measuring the distance in the efficient frontier between an investor hoping for the tangency portfolio with a disappointment profile and the sample efficient frontier is to consider the upper bound factor C_g of $U(\tilde{\mathbf{\Omega}})$ with respect to $U(\hat{\mathbf{\Omega}}_s)$ which is:

$$C_g = 2 \tag{32}$$

Since the tangency portfolio of an investor with the maximum ambiguity is above the sample tangency portfolio, this investor will ask a higher risk premium than the one obtained by the sample tangency portfolio. From Proposition 6, the first derivative of the

[12]Remember that $U(\hat{\mathbf{\Omega}}_s)$ is equal to $\frac{1}{2\gamma}(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)$.

utility function from the covariance matrix with respect to the risk aversion parameter is negative. Therefore, the utility function from our covariance matrix is a decreasing function of the risk aversion parameter. The following expression denotes the second derivative of $U(\tilde{\mathbf{\Omega}})$ with respect to γ:

$$D_{\gamma}^{2}\left[U(\tilde{\mathbf{\Omega}})\right] = \frac{\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_{s}^{-1}\mu}{\gamma^{3}} \tag{33}$$

which is positive. As a result, the risk aversion γ_g of an investor with the maximum ambiguity is obtained by minimizing the utility function:

$$\gamma_{g} = \frac{1}{2}\sqrt{\frac{(\mu^{\mathbf{T}}\mu)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_{s}^{-1}\mu)(N+2)}{T-N-2}} \tag{34}$$

and the resulting expression of the covariance matrix is defined as follows:

$$\tilde{\mathbf{\Omega}}_{g}= \hat{\mathbf{\Omega}}_{s}+\alpha_{g}\mathbf{\Lambda} \tag{35}$$

where α_g is defined as follows:

$$\begin{aligned}\alpha_{g} &= \frac{N^{2}(N+2)(\mu^{\mathbf{T}}\mu)}{4\gamma_{g}^{2}(T-N-2)}\\ &= \frac{N^{2}}{\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_{s}^{-1}\mu}\end{aligned} \tag{36}$$

Notice this result is true when $(N+2) < T$, otherwise γ_g will not exists. The resulting utility function is defined as follows:

$$\begin{aligned}U(\tilde{\mathbf{\Omega}}_{g})&=U(\hat{\mathbf{\Omega}}_{s})+\sqrt{\frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_{s}^{-1}\mu)(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)}}\\ &= 2\sqrt{\frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_{s}^{-1}\mu)(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)}}\end{aligned} \tag{37}$$

and the expected utility function is then defined as:

$$\begin{aligned}E\left[U(\tilde{\mathbf{\Omega}}_{g})\right] &= \frac{T}{T-N-2}U(\hat{\mathbf{\Omega}}_{s}) + \frac{2T-N-2}{\sqrt{N+2}}\sqrt{\frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_{s}^{-1}\mu)}{(\mu^{\mathbf{T}}\mu)(T-N-2)}}\\ &= (3T-N-2)\sqrt{\frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_{s}^{-1}\mu)}{(\mu^{\mathbf{T}}\mu)(T-N-2)(N+2)}}\end{aligned} \tag{38}$$

Assume now an investor computing every weeks his (her) optimal portfolio and buy this allocation one month later. At the end of the week, investor increases the database of returns of one period ($\Delta T = 1$), what corresponds in reality to an increase of $(\Delta T/h)$ where

h denotes the number of weeks in a month. Smaller $\Delta T/h$, and smaller the decrease of the parameter α, and in order to keep constant the investment horizon, higher the parameter h. Notice that, when the decrease of α because of the increase of the database is smaller than the increase of α because of the change in the risk aversion parameter, the sample covariance matrix tends to move away from the covariance matrix. On the contrary, when the decrease of α because of the increase of the database is higher than the increase of α because of the change in the risk aversion parameter, the sample covariance matrix tends to move closer from the covariance matrix. If the risk aversion of investor remains constant from T to $k\Delta T$, thus k denotes the feasible period for the optimal portfolio of investor. The following corrolary states this observation in the case of an investor having the maximum ambiguity reliance on the Sharpe model:

Corollary 1. An investor hoping for the tangency portfolio with the maximum ambiguity reliance on the sample tangency portfolio, should not hold this portfolio more than k/h periods, with h defined as follows:

$$h = \frac{4(T-N-2)^2}{T+k-N-2}\sqrt{\frac{T-N-2}{(\mu^{\mathbf{T}}\mu)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}} \tag{39}$$

where k is the number of period from time T, for which the risk aversion of investor remains constant.

Proof. See appendix 747.

We computed the covariance matrix, for an investor hoping for the tangency portfolio with the maximum ambiguity reliance on the model he (she) uses. In the next sub-section, we focus on an investor with a relative ambiguity between disappointment and elation.

4.2 The Covariance Matrix for Investor with a Relative Ambiguity

We propose now to build among the set of covariance matrices, the one having the less uncertainty between the utility and the expected utility. Let $\tilde{U}$ denotes the uncertainty between the expected utility and the utility of $\tilde{\mathbf{\Omega}}$:

$$\begin{aligned} \tilde{U} &= E\left[U(\tilde{\mathbf{\Omega}})\right] - U(\tilde{\mathbf{\Omega}}) \\ &= \frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{2\gamma(T-N-2)} + \frac{2\gamma T}{(\mu^{\mathbf{T}}\mu)(N+2)} \end{aligned} \tag{40}$$

Let $\tilde{\mathbf{\Omega}}_u$ denotes the covariance matrix minimizing the uncertainty between the expected utility and the utility. By minimizing $\tilde{U}$ with respect to the risk aversion parameter γ, it is straighforward to compute $\tilde{\mathbf{\Omega}}_u$. The first derivative of $\tilde{U}$ with respect to γ is defined as follows:

$$D_\gamma\left(\tilde{U}\right) = -\frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{2\gamma^2(T-N-2)} + \frac{2T}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{41}$$

Since $T < (N+2)$, the first derivative is negative. When $(N+2) < T$, the first derivative is negative if the following condition is true:

$$\left(\frac{2T}{(\mu^{\mathbf{T}}\mu)(N+2)} \leq \frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{2\gamma^2(T-N-2)}\right) \iff \left(\frac{2\gamma(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \leq \frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{2\gamma T}\right) \tag{42}$$

and by adding on the both side of the equality $U(\hat{\mathbf{\Omega}}_s)$ we have the following condition:

$$U(\hat{\mathbf{\Omega}}_s) \leq U(\tilde{\mathbf{\Omega}}_u) \leq \frac{T+N+2}{T}U(\hat{\mathbf{\Omega}}_s) \tag{43}$$

which is true following Proposition 7. We then have a distance in the efficient frontier between an investor having a relative ambiguity and the classical tangency portfolio. Let C_u denotes this distance, we have:

$$C_u = \frac{T+N+2}{T} \tag{44}$$

Since $(N+2) < T$, the factor C_u is lower than the factor C_g obtained in the case of an investor having the maximum ambiguity. As a result, an investor having the maximum ambiguity will ask a higher risk premium than an investor with a relative ambiguity. Following this observation, it seems that the third implication of Kahn and Sarin (1988)[42] for which people are willing to pay for difference in ambiguity is true. Higher the ambiguity reliance on the sample tangency portfolio and higher the risk premium asked by investor.

The second derivative of $\tilde{U}$ with respect to γ is defined as follows:

$$D_\gamma^2\left(\tilde{U}\right) = \frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{\gamma^3(T-N-2)} \tag{45}$$

Since $(N+2) < T$, the second derivative of $\tilde{U}$ with respect to γ is positive. That is, there is a minimum for $\tilde{U}$ obtained by setting $D_\gamma\left(\tilde{U}\right)$ equals zero. We obtain the following risk aversion parameter:

$$\begin{aligned}\gamma_u &= \frac{1}{2}\sqrt{\frac{N+2}{T}}\sqrt{\frac{\left(\mu^{\mathbf{T}}\mu\right)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{(T-N-2)}} \\ &= \sqrt{\frac{N+2}{T}}\gamma_g\end{aligned} \tag{46}$$

By replacing γ_u in the expression of $\tilde{\mathbf{\Omega}}$, we obtain the covariance matrix minimizing the uncertainty between the utility and the expected utility:

$$\tilde{\mathbf{\Omega}}_u = \hat{\mathbf{\Omega}}_s + \alpha_u\mathbf{\Lambda} \tag{47}$$

where α_u is defined as follows:

$$\begin{aligned}\alpha_u &= \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{4\gamma_u^2(T-N-2)} \\ &= \frac{N^2T}{(N+2)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)}\end{aligned} \tag{48}$$

The utility from $\tilde{\mathbf{\Omega}}_u$ is defined as follows:

$$\begin{aligned}U(\tilde{\mathbf{\Omega}}_u) &= U(\hat{\mathbf{\Omega}}) + \sqrt{\frac{(T-N-2)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)}{T(\mu^{\mathbf{T}}\mu)}} \\ &= \frac{T+N+2}{\sqrt{T(N+2)}}\sqrt{\frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)}}\end{aligned} \tag{49}$$

The expected utility of $\tilde{\mathbf{\Omega}}_u$ is then defined as follows:

$$\begin{aligned}E\left[U(\tilde{\mathbf{\Omega}}_u)\right] &= \frac{T}{T-N-2}U(\hat{\mathbf{\Omega}}_s) + \frac{2T-N-2}{\sqrt{T}}\sqrt{\frac{\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu}{(T-N-2)(\mu^{\mathbf{T}}\mu)}} \\ &= \frac{T^2+(N+2)(2T-N-2)}{\sqrt{T(N+2)}}\sqrt{\frac{\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu}{(\mu^{\mathbf{T}}\mu)(N+2)(T-N-2)}}\end{aligned} \tag{50}$$

and the minimum uncertainty $\tilde{U}_{\min}$ of the model is defined as follows:

$$\tilde{U}_{\min} = 2\sqrt{T}\sqrt{\frac{\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu}{(\mu^{\mathbf{T}}\mu)(T-N-2)}} \tag{51}$$

We can establish a result concerning the feasible period for the tangency portfolio with the lower uncertainty between the utility and the expected utility:

Corollary 2. An investor hoping for the tangency portfolio with a relative ambiguity reliance on the sample tangency portfolio, should not hold this portfolio more than k/h periods, with h defined as follows:

$$h = \frac{4(T-N-2)^2}{(N+2)(T+k-N-2)}\sqrt{\frac{T(T-N-2)}{(\mu^{\mathbf{T}}\mu)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)}} \tag{52}$$

Proof. See appendix 848.

We then have the covariance matrix for an investor hoping for the tangency portfolio with a relative ambiguity. The next proposition gives the behavior of the expected utility from this covariance matrix with respect to the expected utility of the covariance matrix characterizing an investor with the maximum ambiguity:

Proposition 8. The expected utility of an investor with a relative ambiguity is higher than the expected utility of an investor having the maximum ambiguity reliance on the sample tangency portfolio.

Proof. See appendix 948.

In the next sub-section, we focus on the second extreme bahavior of an investor facing to uncertainty. This is the case of an investor with the minimum of ambiguity reliance on the sample tangency portfolio.

4.3 The Covariance Matrix for an Investor with the Minimum Ambiguity

For characterizing an investor having an elation profile reliance on the sample tangency portfolio, we minimize the gap between the expected utility and the sample utility function. We then need to compute the risk aversion parameter minimizing the following expression:

$$\begin{aligned}\hat{U} &= E\left[U(\tilde{\mathbf{\Omega}})\right] - U(\hat{\mathbf{\Omega}}_s) \\ &= \frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{2\gamma(T-N-2)} + \frac{2\gamma(2T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)}\end{aligned} \tag{53}$$

Let $\tilde{\mathbf{\Omega}}_t$ denoting the resulting covariance matrix. By minimizing $\hat{U}$ with respect to the risk aversion parameter γ, it is straighforward to compute $\tilde{\mathbf{\Omega}}_t$. The first derivative of $\hat{U}$ with respect to γ is defined as follows:

$$D_\gamma\left(\hat{U}\right) = -\frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{2\gamma^2(T-N-2)} + \frac{2(2T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{54}$$

Since $T < (N+2)$, the first derivative is negative. When $(N+2) < T$, the first derivative is negative if the following condition is true:

$$\left(\frac{2(2T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \leq \frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{2\gamma^2(T-N-2)}\right) \iff \left(\frac{2\gamma(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \leq \frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{2\gamma(2T-N-2)}\right) \tag{55}$$

and by adding on the both side of the inequality $U(\hat{\mathbf{\Omega}}_s)$ we have the following condition:

$$U(\hat{\mathbf{\Omega}}_s) \leq U(\tilde{\mathbf{\Omega}}_t) \leq \frac{2T}{2T-N-2}U(\hat{\mathbf{\Omega}}_s) \tag{56}$$

which is true following Proposition 7. As above, we have a measure of the distance in the efficient frontier between the sample tangency portfolio and the tangency portfolio for an investor with an elation profile. Let C_t this factor of distance, we have:

$$C_t = \frac{2T}{2T-N-2} \tag{57}$$

We introduce the following proposition:

Proposition 9. The distance between the tangency portfolio of an investor with the minimum ambiguity and the sample tangency portfolio, is lower than the distance for an investor with a relative ambiguity, which is lower than the distance for an investor with the maximum ambiguity.

Proof. See appendix 1050.

Following the previous proposition, higher the ambiguity of investor and higher the risk premium asked, which is consistent with the third implication of Kahn and Sarin (1988)[42] for which consumers are willing to pay for differences in ambiguity in choices. Since the risk aversion parameter of the investor with the minimum ambiguity reliance on the sample tangency portfolio is not equal to zero, the first implication of Kahn and Sarin (1988)[42] for which decison makers consider ambiguity when making choice under uncertainty, is true. This observation implies that, in the market, the Sharpe model is always uses by investors with a minimum ambiguity.

The second derivative of $\hat{U}$ with respect to γ is defined as follows:

$$D_\gamma^2\left(\hat{U}\right) = \frac{(\mu^{\mathbf{T}}\hat{\boldsymbol{\Omega}}_s^{-1}\mu)(N+2)}{\gamma^3(T-N-2)} \tag{58}$$

Since $(N+2) < T$, the second derivative of $\hat{U}$ with respect to γ is positive. That is, there is a minimum for $\hat{U}$ obtained by setting $D_\gamma\left(\hat{U}\right)$ equals zero. We obtain the following risk aversion parameter:

$$\begin{aligned}\gamma_t &= \frac{1}{2}\sqrt{\frac{N+2}{2T-N-2}}\sqrt{\frac{\left(\mu^{\mathbf{T}}\mu\right)\,(\mu^{\mathbf{T}}\hat{\boldsymbol{\Omega}}_s^{-1}\mu)(N+2)}{(T-N-2)}} \\ &= \sqrt{\frac{N+2}{2T-N-2}}\gamma_g\end{aligned} \tag{59}$$

Since $(N+2) < T$, we have $\gamma_t < \gamma_g$, as a result γ_t is a relevant risk aversion parameter. By replacing γ_t in the expression of $\tilde{\boldsymbol{\Omega}}$, we obtain the covariance matrix for an investor having the minimum ambiguity:

$$\tilde{\boldsymbol{\Omega}}_t = \hat{\boldsymbol{\Omega}}_s + \alpha_t\boldsymbol{\Lambda} \tag{60}$$

where α_t is defined as follows:

$$\begin{aligned}\alpha_t &= \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{4\gamma_t^2(T-N-2)} \\ &= \frac{N^2(2T-N-2)}{(N+2)(\mu^{\mathbf{T}}\hat{\boldsymbol{\Omega}}_s^{-1}\mu)}\end{aligned} \tag{61}$$

The utility from $\tilde{\boldsymbol{\Omega}}_t$ is defined as follows:

$$\begin{aligned}U(\tilde{\boldsymbol{\Omega}}_t) &= U(\hat{\boldsymbol{\Omega}}) + \sqrt{\frac{(\mu^{\mathbf{T}}\hat{\boldsymbol{\Omega}}_s^{-1}\mu)(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)}}\sqrt{\frac{N+2}{2T-N-2}} \\ &= \frac{2T}{\sqrt{(N+2)(T-N-2)}}\sqrt{\frac{(\mu^{\mathbf{T}}\hat{\boldsymbol{\Omega}}_s^{-1}\mu)(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)}}\end{aligned} \tag{62}$$

The expected utility of $\tilde{\mathbf{\Omega}}_t$ is then defined as follows:

$$
\begin{aligned}
E\left[U(\tilde{\mathbf{\Omega}}_t)\right] &= \frac{T}{T-N-2}U(\hat{\mathbf{\Omega}}_s) + \sqrt{(N+2)(2T-N-2)}\sqrt{\frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)}{(\mu^{\mathbf{T}}\mu)(N+2)(T-N-2)}} \\
&= \frac{(T+N+2)\sqrt{2T-N-2}}{\sqrt{N+2}}\sqrt{\frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)}{(\mu^{\mathbf{T}}\mu)(N+2)(T-N-2)}}
\end{aligned}
\tag{63}
$$

and the minimum uncertainty $\hat{U}_{\min}$ between the expected utility and the utility from the sample covariance matrix is defined as follows:

$$
\hat{U}_{\min} = 2\sqrt{\frac{N+2}{2T-N-2}}\sqrt{\frac{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{(\mu^{\mathbf{T}}\mu)\,(T-N-2)}} \tag{64}
$$

That is, $\tilde{\mathbf{\Omega}}_t$ is the covariance matrix for which the tangency portfolio is almost equals to the sample tangency portfolio, but because of the minimum ambiguity of investors, the two portfolios will never be the same unless the number of historical returns goes to the infinite. We can establish the result concerning the holding period for the corresponding tangency portfolio:

Corollary 3. An investor hoping for the tangency portfolio with the minimum ambiguity reliance on the sample tangency portfolio, should not hold this portfolio more than k/h periods, with h defined as follows:

$$
h = \frac{4(T-N-2)^2}{(N+2)(T+k-N-2)}\sqrt{\frac{(2T-N-2)(T-N-2)}{(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(\mu^{\mathbf{T}}\mu)}} \tag{65}
$$

where k is the number of period from time T, for which the risk aversion of investor remains constant.

Proof. See appendix 1151.

Our model is relevant whether investor hoping for the tangency portfolio and having the minimum ambiguity reliance on the sample mean-variance framework, should have a better expected utility than other investors. Remember, the expected utility is not an optimal criterion when investor facing to uncertainty, however this criterion is consistent with Sharpe (1964)[70]. We state this result in the next proposition:

Proposition 10. The expected utility of an investor having the minimum ambiguity with the sample mean-variance framework is higher than the expected utility of an investor having a relative ambiguity.

Proof. See appendix 1252.

Since the covariance matrix depends on the behavior of investor facing to uncertainty, our model is consistent with the second implication of Khan and Sarin (1988)[42] for which the attitude of investors toward ambiguity varies. The covariance matrix changes with respect to the risk aversion of investor.

Following Proposition 8 and Proposition 10, we know that $\gamma_t < \gamma_u < \gamma_g$, as a result $\alpha_g < \alpha_u < \alpha_t$. However, we know that:

$$\tilde{\mathbf{\Omega}} - \hat{\mathbf{\Omega}}_{\mathbf{s}} = \alpha \mathbf{\Lambda} \tag{66}$$

Therefore the parameter α denotes the under estimation of the risk. An investor having the minimum ambiguity and using the sample covariance matrix, under estimates the risk more than an investor having the maximum ambiguity reliance on the Sharpe model and using the sample covariance matrix. The latter investor being aware about the limit of the sample covariance matrix, will ask a risk premium through C_g higher than the risk premium C_t asked by the former. As a result, in term of risk management, the covariance matrix $\tilde{\mathbf{\Omega}}_t$ is the most efficient. However in term of decision under uncertainty (an *out-of-sample* allocation process), the existence in the market of several investors with several ambiguity profile have some impact in the *out-of-sample* framework. Even when the covariance matrix $\tilde{\mathbf{\Omega}}_t$ taking into account the whole risk is used for this purpose, the existence in the market of investors having a maximum ambiguity has an impact in the *out-of-sample* framework. Therefore, in a decision approach, the covariance matrix to take into account is the one taking into account all the ambiguity in the market. In the next section, we propose to compute the corresponding covariance matrix and to show the inception of ambiguity.

5 The Equilibrium State and the Inception of Ambiguity

The extreme behaviors of investor are characterized by the covariance matrices $\tilde{\mathbf{\Omega}}_g$ for the investor having the maximum ambiguity and $\tilde{\mathbf{\Omega}}_t$ for the investor having the minimum ambiguity reliance on the Sharpe model. In order to explain the inception of ambiguity, we need to build an equilibrium state between the two extreme behaviors.

5.1 The Equilibrium State

Kan and Zhou (2007)[43] state that, under uncertainty, the *out-of-sample* expected loss function generates a loss in the *out-of-sample* performance. They propose to use the global minimum variance portfolio which has some estimation errors not correlated with the estimation errors into the sample tangency portfolio. Since we are able to find an optimal combination between the sample tangency portfolio and the global minimum variance portfolio, we may expect higher *out-of-sample* performance. They show this strategy dominates several approachs among which, the two-fund portfolio rule, the Bayesian approach and the shrinkage estimators under an *i.i.d.* normal multivariate framework.

The covariance matrix $\tilde{\mathbf{\Omega}}_g$ is the one for an investor having the maximum ambiguity, as a result, the weight of the risk-free asset into the resulting tangency portfolio is higher than the one obtained from the llocation of an investor with a minimum ambiguity reliance on the Sharpe model. In the case there is an equilibrium in the market, it is between the two extreme behaviors. Let $\tilde{\mathbf{\Omega}}_{gt}^{\beta}$ the covariance matrix obtained from a linear combination between $\tilde{\mathbf{\Omega}}_g$ and $\tilde{\mathbf{\Omega}}_t$:

$$\tilde{\mathbf{\Omega}}_{gt}^{\beta} = \beta\tilde{\mathbf{\Omega}}_g + (1-\beta)\tilde{\mathbf{\Omega}}_t \tag{67}$$

with $0 < \beta < 1$, we introduce the following proposition:

Proposition 11. The expected utility of the investor having a relative ambiguity is lower than the expected utility of any investor whose the covariance matrix is denoted by $\tilde{\mathbf{\Omega}}_{gt}^{\beta}$:

$$E\left[U(\tilde{\mathbf{\Omega}}_u)\right] \leq E\left[U(\tilde{\mathbf{\Omega}}_{gt}^{\beta})\right] \tag{68}$$

Proof. See appendix 1352.

Therefore, among the infinite ambiguity profile of investors between the extreme behaviors, the expected utility of investor having a relative ambiguity is the lowest. We introduce the following proposition:

Proposition 12. The utility function of the investor having a relative ambiguity is lower than the utility function of any investor whose the covariance matrix is denoted by $\tilde{\mathbf{\Omega}}_{gt}^{\beta}$:

$$U(\tilde{\mathbf{\Omega}}_u) \leq U(\tilde{\mathbf{\Omega}}_{gt}^{\beta}) \tag{69}$$

Proof. See appendix 1453.

Therefore, among the infinite ambiguity profile of investors between the extreme behaviors, the utility of investor having a relative ambiguity is the lowest. This observation means that, between the extreme behaviors, the covariance matrix $\tilde{\mathbf{\Omega}}_u$ is the one for which investor has the maximum level of risk-free asset, we may expect an expected loss lower than usually because of the weight of the risk-free asset in the corresponding portfolio. Therefore the covariance matrix $\tilde{\mathbf{\Omega}}_u$ characterizes the ambiguity profile for which the expected utility can not be improved without increase the level of the expected loss in the *out-of-sample* performance of the tangency portfolio.

The tangency portfolio obtained from $\tilde{\mathbf{\Omega}}_u$ denotes the portfolio from which there is no more possible *trade-off* between the expected utility and the loss in the *out-of-sample* performances. The optimal parameter β is then defined by the following corollary:

Corollary 4. The optimal parameter β^* for the linear combination between investors having the maximum ambiguity and investors having the minimum ambiguity reliance on the Sahrpe model is defined as follows:

$$\beta^* = \sqrt{\frac{N+2}{T}} \frac{\sqrt{2T-N-2} - \sqrt{T}}{\sqrt{2T-N-2} - \sqrt{N+2}} \tag{70}$$

Proof. See appendix 1554.

Now we provided an equilibrium state in the market, we can derive the inception of ambiguity. This is the aim of the next sub-section.

5.2 The Inception of Ambiguity

The next picture shows the shape of the optimal parameter depending on the size of the investment universe:

- Please, insert somewhere here Figure 155 -

From N equals 200, we compute the optimal parameter β^* from T equals $N+3$ to 150,000. We then plot the shape of the optimal parameters obtained with respect to the number of historical returns T. Next, we increment N of 200, the number of assets is now equal to 400, and we compute the optimal parameters and plot them with respect to T. We renew the algorithm fifty times.

Following the picture, it seems that the optimal parameter decreases with respect to the number of historical returns. As a result, higher the number of historical returns in the investment universe and lower the weight of the covariance matrix characterizing the maximum ambiguity reliance on the Sharpe model. Notice that bolder the curve and higher the number of assets considered. Then, higher the number of assets in the investment universe, and lower the decrease of the optimal parameter. This observation means that the weight of the covariance matrix characterizing the maximum ambiguity decreases slowly with respect to the number of historical returns, when the number of assets considered increases. Therefore, the inception of ambiguity of investors comes actually from the size of the investment universe.

It seems than the optimal parameter which characterizes the ambiguity, follows a power law distribution with a paramter almost equal to 1.72. In order to confirm this observation, we show in the next picture, the shape of the optimal parameters (in logarithm) with respect to the ratio N/T (in logarithm).

- Please, insert somewhere here Figure 255 -

For all the combinations, the shapes of the optimal parameter coincide and correspond to a mix of several increasing straight lines. As a result there is some scale effects in the evolution of the optimal parameter with respect to the size of the investment universe characterized by the ratio N/T. The slope of these lines (almost equal to 0.58) provide the parameter of the resulting power law distribution about equals to 1.72.

Kahn and Sarin (1988)[42] propose to consider the entire distribution of the moment of order two in order to take into account the higher moments of the sample returns. We build from our model some covariance matrices for which the entire distribution is taken into account through the parameter α. In the next section, we run an empirical study in the S&P500 universe for showing the consistency of our model.

6 The Empirical Study

We use three databases from the S&P500 universe. Database 1 is constituted of 250 total asset returns in a daily frequency from 01/01/1980 to 03/11/2009. Database 2 is constituted of 250 total asset returns in a weekly frequency from 01/01/1980 to 03/11/2009. Database 3 is constituted of 324 total asset returns in a weekly frequency from 12/22/1988 to 12/18/2008. Returns are all expressed in USD. There is no selection of assets, we only take into account all assets completely available in the financial database between the first and the last dates considered. As a result, there is no completion of returns in our databases. The financial database used for extracting returns is Datastream.

Database 1 and Database 2 are used for showing the consitency of the model. We use Database 3 for the preliminary application of our model to the portfolio optimization because it contains more assets than the two others.

6.1 The Consistency of the Model

In order to show the consistency of the model, we propose to compute the volatility of the risk aversion parameter. The first estimation window for the weekly frequency contains $N+3$ (which is equal to 253) historical returns. For the daily frequency, we consider a number of historical returns corresponding to the same sample dates than the sample window consider for the weekly frequency. We then compute the risk aversion parameter and increase the window size of one day for the daily frequency and one week for the weekly frequency. The next picture shows the dynamic of the risk aversion parameter γ_u:

- Please, insert somewhere here Figure 356 -

Following the picture, the risk aversion parameter decreases as the number of historical returns increases. The decrease is more important with the weekly frequency. However, even there is a decrease of the risk aversion parameter, we notice some variations of its values through time. The next picture shows the variation of the risk aversion parameter and the volatilty of the S&P500 index in a daily frequency:

- Please, insert somewhere here Figure 456 -

We notice the same behavior between the volatility of the risk aversion parameter and the volatility of the market index. The risk aversion parameter characterizes the behavior of an investor facing to the uncertainty. As a result, the variation of the risk aversion parameter should catch all the properties of the volality in the market. When the risk aversion of investors knows some turbulence, the S&P500 index knows some turbulence. When the volatility of the risk aversion of investors is low, the S&P500 index knows a period of low volatility. When considering weekly frequency, we obtain the following picture:

- Please, insert somewhere here Figure 557 -

Although the fit is not as well obvious as for the daily frequency, we observe the same behavior for the two series. The next picture shows how the ambiguity of investors may explained a part of the volatility in the market.

- Please, insert somewhere here Figure 657 -

We have on the same picture the S&P500 index returns and the volatility of the risk aversion in daily frequency on the top, in weekly frequency bottom. It seems that the ambiguity of investors explained a part of the volatility in the market. With the same trend, as the number of historical returns increases, the volatility of the risk aversion decreases, which means that its the impact of the ambiguity in the market volatility decreases. An investor having the maximum ambiguity, will therefore ask a higher risk premium and his (her) will support only the part of the market volatility not explained by his (her) ambiguity. As a result, the use by this investor of the sample covariance matrix yield less errors in the estimation of the risk. In the same manner, an investor having the minimum ambiguity is not aware of this risk. The investor will ask a lower risk premium and the use of the sample covariance matrix yield much errors in the estimation of the risk.

As seen above, the approach based on the risk management requires the use of the covariance matrix characterizing well the risk depending on the ambiguity of investor. Concerning an allocation process under uncertainty, the existence of several profile of investors in relation to ambiguity, has for consequence the need to aggregate the ambiguity in the market. Using Database 3, in the next sub-section we make a preliminary application to the portfolio optimization.

6.2 Preliminary Application to the Portfolio Optimization

From Jagannathan and Ma (2003)[35], we know that the imposition of a shortsale constraint in the portfolio optimization program improves the performance of the sample estimator of the covariance matrix, a fortiori the performance of the sample tangency portfolio. We only take into account in our empirical study the case where there is a shortsale constraint.

Usually, academic literature uses an *out-of-sample* approach with a rolling estimation window in order to remains constant the number of historical returns. We know that the sample estimators of the expected returns and the covariance matrix perform better with the increase of the number of historical returns. In order to take into account the effect of the sample size neglect through time, we consider a more conservative approach by using an increasing estimation window for the *out-of-sample* portfolio.

However, as noticed above, our model allows to understand the dynamic between the sample covariance matrix and the true one. The most important are not the values, but their variation through time. In order to use our covariance matrices in an allocation process, we need to apply a scale parameter on parameter α. This is the aim of the next sub-section.

6.2.1 The Scale Parameter

We need a scale parameter for calibrating our covariance matrices at the same order than elements of the sample covariance matrix. Notice that, among our three covariance matrices, there is a common factor having the following expression:

$$\frac{N^2}{\mu^{\mathbf{T}} \hat{\mathbf{\Omega}}_s^{-1} \mu} \mathbf{\Lambda}$$

We propose to scale the common factor among the three covariance matrices with the sample one. Notice that by this approach, we cancel the parameter α_g in the expression of the covariance matrix $\tilde{\mathbf{\Omega}}_g$, and reduce its elements on the same order than the elements of the sample covariance matrix. By canceling the parameter α_g, we assume that the entire distribution of the order two moment characterizes the most ambiguous profile reliance on the Sharpe model. For the two other covariance matrices, it remains a weighting parameter depending only of the size of the investment universe. We then propose the following scale parameter denoted by ζ:

$$\zeta = \frac{N^2}{\mu^{\mathbf{T}} \hat{\mathbf{\Omega}}_s^{-1} \mu} \sum_{i,j} \frac{\mathbf{\Lambda}_{ij}}{\hat{\mathbf{\Omega}}_{s_{ij}}} \tag{71}$$

We divide the weighting parameter α by the scale parameter ζ and obtain covariance matrices for a practical use.

6.2.2 The Portfolio Allocation Process

The optimization program when there is a shortsale constraint is given by:

$$\begin{cases} \underset{(x_p)}{Max} \dfrac{x_p^{'}\hat{\mu}_s}{x_p^{'}\mathbf{\Omega} x_p} \\ s.t\ x_p^{'}\mathbf{1} = 1 \\ x_{p_i} \geq 0,\ i = 1, ..., N \end{cases} \tag{72}$$

where $\hat{\mu}_s$ denotes the sample expected returns and $\mathbf{\Omega}$ one the covariance matrices among $\hat{\mathbf{\Omega}}_s$, $\tilde{\mathbf{\Omega}}_g$, $\tilde{\mathbf{\Omega}}_t$ and $\tilde{\mathbf{\Omega}}_u$. We consider the first estimation window from 12/22/1988 to 03/16/1995, we have 324 assets with 326 historical returns. We then compute the optimal tangency portfolio, and buy it. One week later, we increase the number of historical returns of one period, we thus have a new estimation window from 12/15/1988 to 03/23/1995 that means 324 assets and 327 historical returns. We compute the new optimal tangency portfolio and rebalance the allocation by selling or buying assets from the last allocation in order to match with the new allocation. We perform the algorithm until the end date. We finally obtain an *out-of-sample* tangency portfolio from 03/23/1995 to 12/18/2008. We obtain four tangency portfolios from $\hat{\mathbf{\Omega}}_s$, $\tilde{\mathbf{\Omega}}_g$, $\tilde{\mathbf{\Omega}}_t$ and $\tilde{\mathbf{\Omega}}_u$ and compute their corresponding Sharpe ratios. Without loss of generality we do not consider the risk-free asset in our database.

Next, we consider one thousand possible linear combinations between the tangency portfolios obtained from $\tilde{\mathbf{\Omega}}_g$, $\tilde{\mathbf{\Omega}}_t$ and we compute their corresponding Sharpe ratios. The following picture shows the Sharpe ratio of all the tangency portfolios considered:

- Please, insert somewhere here Figure 758 -

6.2.3 Comments

From Figure 7, it appears that the Sharpe ratio[13] of the sample tangency portfolio equals to 0.94, is higher than the one of the tangency portfolio obtained from $\tilde{\mathbf{\Omega}}_g$ (equals to 0.9). However it remains lower than the Sharpe ratio obtained from the covariance matrix $\tilde{\mathbf{\Omega}}_t$ (equals to 0.98). The covariance matrix $\tilde{\mathbf{\Omega}}_u$ is the one characterizing the equilibrium state between the most and the less ambiguous investors. The empirical result seems testify our theoretical result; the Sharpe ratio of the tangency portfolio obtained from the covariance matrix $\tilde{\mathbf{\Omega}}_u$ (equals to 1.03) is higher than the Sharpe ratio obtained from the covariance matrix having the highest expected utility and then, a fortiori than the Sharpe ratio of the sample tangency portfolio.

We know that the covariance matrix $\tilde{\mathbf{\Omega}}_u$ is actually a linear combination between the $\tilde{\mathbf{\Omega}}_g$ and $\tilde{\mathbf{\Omega}}_t$ with the optimal parameter β^*. In order to see if there is a naive approach

[13] There is no cash considered in our expression of the Sharpe ratio. This is an annualized Sharpe ratio.

allowing to build a tangency portfolio with a most higher Sharpe ratio than the one obtained from the covariance matrix $\tilde{\mathbf{\Omega}}_u$, we also represent on Figure 6, the Sharpe ratios of one thousand tangency portfolios obtained from several linear combinations between $\tilde{\mathbf{\Omega}}_g$ and $\tilde{\mathbf{\Omega}}_t$. From Figure 7, it seems that there is some combinations having a lower and some having a higher Sharpe ratio than the sample tangency portfolio. We also observe some combinations having a higher Sharpe ratio than the tangency portfolio obtained from $\tilde{\mathbf{\Omega}}_t$. However, none of these combinations do not have a higher Sharpe ratio than the one obtained from the covariance matrix $\tilde{\mathbf{\Omega}}_u$ characterizing the equilibrium state.

7 Conclusion

In this paper, we develop a model for understanding the impact of the sample size neglect of investor when he (she) hopes for the tangency portfolio. From the literature dealing with parameter uncertainty, we compute the expected loss function of the sample covariance matrix under a wrong hypothesis. By inverting the expected loss function of the sample covariance matrix and taking its opposite, we track the dynamic between the sample covariance matrix and the true covariance matrix and highlight the impact of the ambiguity in the measure of the volatility of the portfolio. We then distinguish a covariance matrix for the risk management issue and the covariance matrix for an allocation decision under uncertainty.

The impact of the sample size neglect has been shown. It seems that an investor not aware on the sample size of the investment universe will ask a lower risk premium than another. Therefore, by using the sample covariance matrix, this investor under estimates the risk of his (her) portfolio. The model allows to build a covariance matrix allowing to well measure the risk.

An equilibrium state has been built in order to deal with the problem of allocation decision under uncertainty. The equilibrium allows to show that the ambiguity of investors comes actually from the sample size of the investment universe and follows a power law distribution with respect to the ratio between the number of assets and the number of historical returns, with a parameter almost aquals to 1.72. A preliminary result on the american market seems confirm the relevance of the covariance matrix derived from the equilibrium state.

The model is consistent with the litarature dealing with the ambiguity, and with the non-normal character of asset returns in markets. It may be helpful for explaining some financial puzzles as the risk premium puzzle or the volatility puzzle.

References

[1] Barberis N. and R. Thaler, 2003, "A Survey of Behavioral Finance", *in G.M. Constantinides, M. Harris and R. Stulz eds., Handbook of the Economics of Finance*, 71 pages.

[2] Barry C.B., 1974, "Portfolio Analysis under Uncertain Means, Variances, and Covariances", *Journal of Finance 29*, 515-522.

[3] Bawa V.S., S. Brown and R. Klein, 1979, "Estimation Risk and Optimal Portfolio Choice", *North Holland, Amsterdam.*

[4] Best M.J. and R.R. Grauer, 1991, "On the Sensitivity of Mean-Variance-efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results", *Review of Financial Studies 4*, 315-342.

[5] Black F., M.C. Jensen and M. Scholes, 1972, "The CAPM: Some Empirical Tests", *in Jensen Studies in the Theory of Capital Markets, New York, Praeger Publishers*, 79-121.

[6] Brandt M., 2004, "Portfolio Choice Problems", *forthcoming in Y. Ait-Sahalia and L.P. Hansen eds., Handbook of Financial Econometrics, Elsevier Science: Amsterdam*, 75 pages.

[7] Brown S.J., 1978, "The Portfolio Choice Problem: Comparison of Certainty Equivalence and Optimal Bayes Portfolios", *in Communications in Statistics-Simulation and Computation 7*, 321-334.

[8] Chauveau T. and N. Nalpas, 2009, "A Theory of Disappointment", *ESC-Toulouse Working Papers*, 26 pages.

[9] Chopra V.K. and W.T. Ziemba, 1993, "The Effects of Errors in Means, Variances and Covariances on Optimal Portfolio Choice", *Journal of Portfolio Management 19*, 6-11.

[10] Chopra V.K., 1993, “Improving Optimization”, *Journal of Investing 8*, 51-59.

[11] Conlon T., H.J. Ruskin and M. Crane, 2008, “Random Matrix Theory and Fund of Funds Portfolio Optimisation”, *Dublin City University, working paper*, 16 pages.

[12] Cont R., 1998, “Empirical Properties of Assets Returns: Stylized Facts and Statistical Issues”, *Quantitative Finance 1*, 223-236.

[13] DeMiguel V., L. Garlappi, F.J. Nogales and R. Uppal, 2007, “A Generalized Approach to Portfolio Optimization: Improving Performance By Constraining Portfolio Norms”, *forthcoming in Management Science*, 36 pages.

[14] DeMiguel V., L. Garlappi, and R. Uppal, 2007, “Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?”, *forthcoming in Review of Financial Studies*, 39 pages.

[15] DeMiguel V. and F.J. Nogales, 2007, “Portfolio Selection with Robust Estimation of Risk”, *Forthcoming in Operations Research, manuscript OPRE-2007-02-106.R3*, 39 pages.

[16] Disatnik D. and S. Benninga, 2007, “Shrinking Covariance Matrix - The Simpler is Better”, *Journal of Portfolio Management 33*, 56-63.

[17] Edelman A., 1989, “Eigenvalues and Condition Numbers of Random Matrices”, *PhD Thesis, Department of Mathematics, Massachussetts Institute of Technology, Cambridge, MA*, 105 pages.

[18] Ellsberg D., 1961, “Risk, Ambiguity, and the Savage Axioms”, *Quarterly Journal of Economics 75*, 643-669.

[19] Elton E. and M. Gruber, 1973, “Estimating the Dependence Structure of Share Prices”, *Journal of Finance 28*, 1203-1232.

[20] Fama E.F. and K.R French, 1993, "Common Risk Factors in the Returns of Stock and Bonds", *Journal of Financial Economics 33*, 3-56.

[21] Fama E.F. and K.R French, 2004, "The Capital Asset Pricing Model: Theory and Evidence", *The Journal of Economic Perspectives 18*, 25-46.

[22] Fan J., Y. Fan and J. Jiang, 2007, "Dynamic Integration of Time- and State-Domain Methods for Volatility Estimation", *Journal of American Statistical Association 102*, 618-631.

[23] Frost P. and J. Savarino, 1988, "For Better Performance Constrain Portfolio Weights", *Journal of Portfolio Management 15*, 29-34.

[24] Garlappi L., R. Uppal and T. Wang, 2007, "Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach", *Review of Financial Studies 20*, 41-81.

[25] Gilboa I. and D. Schmeidler, 1989, "Maxmin Expected Utility with Non-unique Prior", *Journal of Mathematical Economics 18*, 141-153.

[26] Goldfarb D. and G. Iyengar, 2003, "Robust Portfolio Selection Problems", *Mathematics of Operations Research 28*, 1-38.

[27] Gul F.,1991, "A Theory of Disappointment in Decision Making Under Uncertainty", *Econometrica 59*, 667-686.

[28] Hansen L.P. and T.J. Sargent, 2001, "Robust Control and Model Uncertainty", *American Economic Association 91(2)*, 60-66.

[29] Hosking J., J. Wallis and J. Wood, 1985, "Estimation of the Generalized Extreme-value Distribution by the Method of Probability-weighted Moments", *Technometrics 27*, 251-261.

[30] Hosking J., 1986, "The Theory of Probability Weighted Moments", *Research Report RC12210, IBM Research Division*, 15 pages.

[31] Hosking J. and J. Wallis, 1987, "Parameter and Quantile Estimation for the Generalized Pareto Distribution", *Technometrics 29*, 339-349.

[32] Hosking J., 1989, "Some Theoretical Results Concerning L-Moments", *Research Report RC14492, IBM Research Division*, 13 pages.

[33] Hosking J., 1990, "L-moments: Analysis and Estimation", *Journal of the Royal Statistical Society Series B 52*, 105-124.

[34] Jagannathan R. and T. Ma, 2000, "Three Methods for Improving the Precision in Covariance Matrix Estimation", *Manuscript, Kellogg School of Management, Northwestern University.*

[35] Jagannathan R. and T. Ma, 2003, "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps", *Journal of Finance 54*, 1651-1683.

[36] Jobson J.D. and R. Korkie, 1980, "Estimation for Markowitz Efficient Portfolios", *Journal of the American Statistical Association 75*, 544-554.

[37] Jobson J.D., R. Korkie and V. Ratti, 1979, "Improved Estimation for Markowitz Portfolios Using James-Stein Type Estimators", *American Statistical Association 41*, 279-292.

[38] Jorion P., 1985, "International Portfolio Diversification with Estimation Risk", *Journal of Business 58*, 259-278.

[39] Jorion P., 1986, "Bayes-Stein Estimation for Portfolio Analysis", *Journal of Financial and Quantitative Analysis 21*, 279-292.

[40] Kahneman D. and A. Tversky, 1974, "Judgment Under Uncertainty: Heuristics and Biases", *Science 185*, 1124-1131.

[41] Kahneman D. and A. Tversky, 1979, "Prospect Theory: An analysis of Decision Under Risk", *Econometrica 47*, 263-291.

[42] Kahn B.E. and R.K. Sarin, 1988, "Modeling Ambiguity in Decisions Under Uncertainty", *Journal of Consumer Research 15*, 265-272.

[43] Kan R. and G. Zhou, 2007, "Optimal Portfolio Choice with Parameter Uncertainty", *Journal of Financial and Quantitative Analysis 42*, 621-656.

[44] Kondor I., S. Pafka and G. Nagy, 2002, "Noise Sensitivity of Portfolio Selection Under Various Risk Measures", *Journal of Banking and Finance 31*, 1545-1573.

[45] Laloux L., P. Cizeau, J.-P. Bouchaud and M. Potters, 1999, "Noise Dressing of Financial Correlation Matrices", *Physical Review Letter 83*, 14-69.

[46] Ledoit O. and M. Wolf, 2003, "Improved Estimation of the Covariance Matrix of Stock Returns With an Application to Portfolio Selection", *Journal of Empirical Finance 10*, 603-621.

[47] Ledoit O. and M. Wolf, 2004-a, "A Well-Conditioned Estimator for Large Dimensional Covariance Matrices", *Journal of Multivariate Analysis 88*, 365-411.

[48] Ledoit O. and M. Wolf, 2004-b, "Honey, I Shrunk the Sample Covariance Matrix", *Journal of Portfolio Management 31*, 1-22.

[49] Litterman R., 2003, "Modern Investment Management: An Equilibrium Approach", *Wiley, New York.*

[50] Loomes G. and R. Sugden, 1982, "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty", *The Economic Journal 92*, 805-824.

[51] MacKinlay A.C. and L. Pástor, 2000, "Asset Pricing Models: Implications for Expected Returns and Portfolio Selection", *Review of Financial Studies 13*, 883-916.

[52] Mandelbrot B., 1963, "The Variation of Certain Speculative Prices", *Journal of Business XXXVI*, 392-317.

[53] Markowitz H., 1952, "Mean-Variance Analysis in Portfolio Choice and Capital Markets", *Journal of Finance 7* , 77-91.

[54] Merton R., 1980, "On Estimating the Expected Return on the Market: An Exploratory Investigation", *Journal of Financial Economics 8(4)*, 323-361.

[55] Meucci A., 2005, "Risk and Asset Allocation ", *New York, Springer-Verlag.*

[56] Michaud R.O., 1989, "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?", *Financial Analysts Journal 45*, 31-42.

[57] Michaud R.O., 1998, "Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization", *Oxford University Press.*

[58] Michaud R.O. and R. Michaud, 2008, "Estimation Error and Portfolio Optimization: A Resampling Solution", *Journal of Investment Management 6(1)*, 8-28.

[59] Muirhead R.J., 1982, "Aspects of Multivariate Statistical Theory", *New-York, Wiley.*

[60] Pafka S. and I. Kondor, 2004, "Estimated Correlation Matrices and Portfolio Optimization", *Physica A 343*, 623-634.

[61] Pástor L., 2000, "Portfolio Selection and Asset Pricing Models", *Journal of Finance 55*, 179-223.

[62] Pástor L. and R.F. Stambaugh, 2000, "Comparing Asset Pricing Models: An Investment Perspective", *Journal of Financial Economics 56*, 335-381.

[63] Plerou V., P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, T. Guhr and H.E. Stanley, 2002, "A Random Matrix Approach to Cross-Correlations in Financial Data", *Physical Review E 65*, 19 pages.

[64] Potters M., J.-P. Bouchaud and L. Laloux, 2005, "Financial Applications of Random Matrix Theory: Old Laces and New Pieces", *Acta Physica Polonica B 36(9)*, 2767-2784.

[65] Rabin M., 2002, "Inference by Believers in the Law of Small Numbers", *Quarterly Journal of Economics 117*, 775-816.

[66] Rousseeuw P.J. and K. Van Driessen, (1999), "A Fast Algorithm for the Minimum Covariance Determinant Estimator", *Technometrics 41*, 212-223.

[67] Savage L., 1964, "The Foundations of Statistics", *Wiley, New York*.

[68] Serfling R. and P. Xiao, 2007, "A Contribution to Multivariate L-Moments: L-Comoment Matrices", *Journal of Multivariate Analysis 98*, 1765-1781.

[69] Sharpe W., 1963, "A Simplified Model for Portfolio Analysis", *Journal of Management Science 9*, 277-293.

[70] Sharpe W., 1964, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk", *Journal of Finance 19(3)*, 425-442.

[71] Sillitto G., 1951, "Interrelations between Certain Linear Systematic Statistics of Sample from any Continuous Population", *Biometrika 38*, 377-382.

[72] Tobin J., 1958, "Liquidity Preferences as Behavior Towards Risk", *Review of Economics Studies 25*, 65-86.

[73] Von-Neumann J. and O. Morgenstern, 1944, "Theory of Games and Economic Behavior", *Princeton University Press*, 776 pages.

[74] Yanou G., 2008, "Extension of Random Matrix Theory to the L-moments for Robust Portfolio Allocation", *Univesity of Paris-1, Working Paper*, 40 pages.

A Appendix

A.1 Proofs of Proposition and Corollary

A.1.1 Proof of Proposition 1.

Proposition 1. If the identity matrix is the covariance matrix, the expected loss function of the sample covariance matrix ρ is defined as follows:

$$\rho = -\frac{(\mu^{\mathbf{T}}\mu)(N+2)}{2\gamma(T-N-2)}$$

Proof.
The expected loss function of the sample covariance matrix is defined as follows:

$$\begin{aligned}\rho_{Id}\left(\mathbf{I},\hat{\mathbf{\Omega}}\right) &= U(\mathbf{I}) - E\left[U(\hat{\mathbf{\Omega}}_s)\right] \\ &= \frac{1}{2\gamma}\mu^{\mathbf{T}}\mu - \frac{1}{2\gamma}\mu^{\mathbf{T}}E(\hat{\mathbf{\Omega}}_s^{-1})\mu\end{aligned} \tag{A.1.1}$$

The sample covariance matrix $\hat{\mathbf{\Omega}}_s$ has the following distribution:

$$\hat{\mathbf{\Omega}}_s \backsim W_N\left(T-1,\mathbf{I}\right)/T \tag{A.1.2}$$

where $W_N\left(T-1,\mathbf{I}\right)$ denotes a Wishart distribution with $T-1$ degree of freedom. Following Muirhead (1982) we know under the null hypothesis that:

$$E(\hat{\mathbf{\Omega}}_s^{-1}) = \frac{T}{T-N-2}\mathbf{I}^{-1} \tag{A.1.3}$$

We obtain a new expression of $\rho\left(\mathbf{I},\hat{\mathbf{\Omega}}_s\right)$:

$$\begin{aligned}\rho\left(\mathbf{I},\hat{\mathbf{\Omega}}_s\right) &= \frac{1}{2\gamma}\mu^{\mathbf{T}}\mu - \frac{1}{2\gamma}\frac{T}{T-N-2}\mu^{\mathbf{T}}\mu \\ &= \frac{\mu^{\mathbf{T}}\mu}{2\gamma}\left(1-\frac{T}{T-N-2}\right)\end{aligned} \tag{A.1.4}$$

and the development of (A.1.4) gives the following expression in terms of distance:

$$\rho_{Id}\left(\mathbf{I},\hat{\mathbf{\Omega}}_s\right) = -\frac{(\mu^{\mathbf{T}}\mu)(N+2)}{2\gamma(T-N-2)} \tag{A.1.5}$$

we then obtain the result.

■

A.1.2 Proof of Proposition 2.

Proposition 2. A candidate $\tilde{\mathbf{\Omega}}$ for the covariance matrix is defined as follows:

$$\tilde{\mathbf{\Omega}} = \hat{\mathbf{\Omega}}_{\mathbf{s}} + \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{4\gamma^2(T-N-2)}\mathbf{\Lambda}$$

where $\mathbf{\Lambda}$ the matrix of size $N \times N$ obtained from the vector of expected returns μ:

$$\mathbf{\Lambda} = \mu \otimes \mu^{\mathbf{T}}$$

with the sign $\otimes$ denoting the kronecker product.

Proof.

If ρ is the asymptotical expected loss function of the sample covariance matrix when we assume the null hypothesis which is actually wrong, $1/\rho$ may be viewed as a measure of the maximum noise between the sample covariance matrix and the covariance matrix:

$$U(\tilde{\mathbf{\Omega}}) - U(\hat{\mathbf{\Omega}}_s) \leq \varrho \tag{A.2.1}$$

where ϱ is equal to $1/\rho$. We assume a conservative point of view:

$$U(\tilde{\mathbf{\Omega}}) - U(\hat{\mathbf{\Omega}}_s) = \frac{2\gamma(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{A.2.2}$$

We then obtain:

$$\mu^{\mathbf{T}}\tilde{\mathbf{\Omega}}^{-1}\mu = \mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu + \frac{4\gamma^2}{\mu^{\mathbf{T}}\mu}\frac{T-N-2}{N+2} \tag{A.2.3}$$

A candidate for which the condition above is true is defined as follows:

$$\tilde{\mathbf{\Omega}} = \hat{\mathbf{\Omega}}_s + \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{4\gamma^2(T-N-2)}\mathbf{\Lambda} \tag{A.2.4}$$

where:

$$\mathbf{\Lambda} = \mu \otimes \mu^{\mathbf{T}}$$

To show this result, we have to compute the utility function of the candidate for the covariance matrix:

$$\begin{aligned}
U(\tilde{\mathbf{\Omega}}) &= \frac{1}{2\gamma}\mu^{\mathbf{T}}\tilde{\mathbf{\Omega}}^{-1}\mu \\
&= \frac{1}{2\gamma}\mu^{\mathbf{T}}\left[\hat{\mathbf{\Omega}}_s + \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{4\gamma^2(T-N-2)}\mathbf{\Lambda}\right]^{-1}\mu \\
&= U(\hat{\mathbf{\Omega}}_s) + \frac{1}{2\gamma}\mu^{\mathbf{T}}\left[\frac{4\gamma^2(T-N-2)}{N^2(N+2)(\mu^{\mathbf{T}}\mu)}\mathbf{\Lambda}^{-1}\right]\mu \\
&= U(\hat{\mathbf{\Omega}}_s) + \frac{2\gamma(T-N-2)}{N^2(N+2)(\mu^{\mathbf{T}}\mu)}\mu^{\mathbf{T}}\mathbf{\Lambda}^{-1}\mu
\end{aligned} \tag{A.2.5}$$

We know that the matrix $\mathbf{\Lambda}$ is equal to the kronecker product between μ and $\mu^{\mathbf{T}}$. Therefore:

$$\begin{aligned}\mu^{\mathbf{T}}\mathbf{\Lambda}^{-1}\mu &= \mu^{\mathbf{T}}\left(\mu\otimes\mu^{\mathbf{T}}\right)^{-1}\mu \\ &= \mu^{\mathbf{T}}\left(\mu^{-1}\otimes\mu^{-1^{\mathbf{T}}}\right)\mu\end{aligned} \tag{A.2.6}$$

Remember μ is the expected returns vector of size $N\times 1$. Assume we have two assets in the investment universe ($N=2$), we then have:

$$\begin{aligned}\mu^{\mathbf{T}}\mathbf{\Lambda}^{-1}\mu &= (\mu_1,\mu_2)\left(\begin{pmatrix}\frac{1}{\mu_1}\\ \frac{1}{\mu_2}\end{pmatrix}\otimes\left(\frac{1}{\mu_1},\frac{1}{\mu_2}\right)\right)\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} \\ &= (\mu_1,\mu_2)\begin{pmatrix}\frac{1}{\mu_1^2} & \frac{1}{\mu_1\mu_2}\\ \frac{1}{\mu_1\mu_2} & \frac{1}{\mu_2^2}\end{pmatrix}\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} \\ &= \left(\frac{2}{\mu_1},\frac{2}{\mu_2}\right)\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} \\ &= 2+2 = 2\times 2 = 2^2\end{aligned} \tag{A.2.7}$$

We assume now this result is true when there is N assets in the investment universe. We then have:

$$\mu^{\mathbf{T}}\mathbf{\Lambda}^{-1}\mu = \underbrace{N+N+N+...+N}_{N\text{ times}} = N\times N = N^2 \tag{A.2.8}$$

By recurrence, we show it remains true when we consider $N+1$ assets in the investment universe. We then have $\mu=(\mu_i)_{i=1,...,N+1}$ and $\mathbf{\Lambda}=\mu^{-1}\otimes\mu^{-1^{\mathbf{T}}}$ of size $(N+1)\times(N+1)$:

$$\begin{aligned}\mu^{\mathbf{T}}\mathbf{\Lambda}^{-1}\mu &= (\mu_i)^T_{i=1,...,N+1}\left[(\mu_i)^{-1}_{i=1,...,N+1}\otimes(\mu_i)^{-1^{\mathbf{T}}}_{i=1,...,N+1}\right](\mu_i)_{i=1,...,N+1} \\ &= \left[(\mu_i)^T_{i=1,...,N}\left[(\mu_i)^{-1}_{i=1,...,N}\otimes(\mu_i)^{-1^{\mathbf{T}}}_{i=1,...,N}\right](\mu_i)_{i=1,...,N}\right] \\ &\quad + \left[(\mu_i)^T_{i=1,...,N+1}\begin{pmatrix}\mathbf{0}_{(i,j)_{i,j=1,...,N}} & \mathbf{\Lambda}_{(i,N+1)_{i=1,...,N}}\\ \mathbf{\Lambda}_{(N+1,j)_{j=1,...,N}} & \mathbf{\Lambda}_{(N+1,N+1)}\end{pmatrix}(\mu_i)_{i=1,...,N+1}\right]\end{aligned} \tag{A.2.9}$$

As a result, we have the following expression:

$$\begin{aligned}\mu^{\mathbf{T}}\mathbf{\Lambda}^{-1}\mu &= N^2+\left[\left((\mu_i)^{-1^T}_{i=1,...,N},\frac{N+1}{\mu_{N+1}}\right)(\mu_i)_{i=1,...,N+1}\right] \\ &= N^2+\underbrace{1+1+1+...+1}_{N\text{ times}}+(N+1) \\ &= \underbrace{N+N+N+...+N}_{N\text{ times}}+\underbrace{1+1+1+...+1}_{N\text{ times}}+(N+1) \\ &= \underbrace{(N+1)+(N+1)+(N+1)+...+(N+1)}_{N+1\text{ times}} \\ &= (N+1)\times(N+1) = (N+1)^2\end{aligned} \tag{A.2.10}$$

Therefore, if we assume N assets in the investment universe, we have:

$$\begin{aligned}U(\tilde{\mathbf{\Omega}}) &= \frac{1}{2\gamma}\mu^{\mathbf{T}}\tilde{\mathbf{\Omega}}^{-1}\mu \\ &= U(\hat{\mathbf{\Omega}}_s) + \frac{2\gamma(T-N-2)}{N^2(N+2)(\mu^{\mathbf{T}}\mu)}\left(\mu^{\mathbf{T}}\mathbf{\Lambda}^{-1}\mu\right) \\ &= U(\hat{\mathbf{\Omega}}_s) + \frac{2\gamma N^2(T-N-2)}{N^2(N+2)(\mu^{\mathbf{T}}\mu)} \\ &= U(\hat{\mathbf{\Omega}}_s) + \frac{2\gamma(T-N-2)}{(N+2)(\mu^{\mathbf{T}}\mu)}\end{aligned} \tag{A.2.11}$$

Since from the candidate, we obtain the same level of noise between the sample covariance matrix and the true one, the candidate is actually the covariance matrix.

■

A.1.3 Proof of Proposition 3.

Proposition 3. The covariance matrix obtained as a sum of the sample covariance matrix and a weighted matrix of expected returns is a defined positive matrix.

Proof.

We first consider the case where $N = 2$. We have the following expression of the covariance matrix:

$$\tilde{\mathbf{\Omega}}_2 = \begin{pmatrix} E\left(X_1^2\right) - \mu_1^2 & E\left(X_1X_2\right) - \mu_1\mu_2 \\ E\left(X_1X_2\right) - \mu_1\mu_2 & E\left(X_2^2\right) - \mu_2^2 \end{pmatrix} + \alpha_2 \begin{pmatrix} \mu_1^2 & \mu_1\mu_2 \\ \mu_1\mu_2 & \mu_2^2 \end{pmatrix} \tag{A.3.1}$$

where μ_i, α_2 and $\tilde{\mathbf{\Omega}}_2$ denote respectively the expected returns of the random variables X_i $(i = 1, \quad ,2)$, the parameter α when $N = 2$ and the covariance matrix of size 2×2. We know that the sample covariance matrix is a defined positive matrix, as a result, we have:

$$\left|\hat{\mathbf{\Omega}}_s\right| = \left[E\left(X_1^2\right)E\left(X_2^2\right) - E\left(X_1X_2\right)^2\right] - \left[\mu_2^2 E\left(X_1^2\right) + \mu_1^2 E\left(X_2^2\right) - 2E\left(X_1X_2\right)\mu_1\mu_2\right] \tag{A.3.2}$$

where $|\cdot|$ denotes the determinant. Since the sample covariance matrix is a defined positive matrix, we have:

$$E\left(X_1^2\right)E\left(X_2^2\right) - E\left(X_1X_2\right)^2 > \mu_2^2 E\left(X_1^2\right) + \mu_1^2 E\left(X_2^2\right) - 2E\left(X_1X_2\right)\mu_1\mu_2 \tag{A.3.3}$$

The expression of the determinant of the covariance matrix is the following:

$$\begin{aligned}\left|\tilde{\mathbf{\Omega}}_2\right| &= \left|\hat{\mathbf{\Omega}}_s\right| + \alpha_2\left[\mu_2^2 E\left(X_1^2\right) + \mu_1^2 E\left(X_2^2\right) - 2E\left(X_1X_2\right)\mu_1\mu_2\right] \\ &\simeq \left|\hat{\mathbf{\Omega}}_s\right| + \alpha_2\left(\mu_2\sqrt{E\left(X_1^2\right)} + \mu_1\sqrt{E\left(X_2^2\right)}\right)\end{aligned} \tag{A.3.4}$$

Since the parameter α_2 is always positive, as a result we have:

$$\left|\tilde{\mathbf{\Omega}}\right| > 0 \tag{A.3.5}$$

Now, we assume that $\left|\tilde{\mathbf{\Omega}}_N\right|$ is positive for N assets, and show the result remains true for $N+1$ assets. The size of the covariance matrix in this case $N+1 \times N+1$. The determinant of this matrix may be obtained by the determinant of some sub-matrix of size $N \times N$. We have:

$$\left|\tilde{\mathbf{\Omega}}_{N+1}\right| = \sum_{i=1}^{N+1} (-1)^{i+j} \left|\tilde{\mathbf{\Omega}}_N^j\right|_{j=1\ldots N+1} \tag{A.3.6}$$

We just have to consider for computing $(-1)^{i+j}$, the diagonal elements of $\tilde{\mathbf{\Omega}}_{N+1}$, for which $(-1)^{i+j}$ is always equal to one because the power coefficients are always odd. We also know that $\left|\tilde{\mathbf{\Omega}}_N\right|$ is always positive. Therefore $\left|\tilde{\mathbf{\Omega}}_{N+1}\right|$ is a sum of positive elements, as a result remains positive.

■

A.1.4 Proof of Proposition 4.

Proposition 4. As the number of historical returns increases from T to $k\Delta T$, the sample covariance matrix goes toward the true covariance matrix from the point $T + k\Delta T$ if the following condition is true:

$$\langle \Delta\gamma_k \rangle \leq \frac{k\Delta T}{h}$$

with:

$$h = \frac{2(\gamma + \Delta\gamma_k)^2 (T - N - 2)^2}{\gamma^3 (T + k\Delta T - N - 2)}$$

where $\langle\cdot\rangle$ denotes the absolute value, $\Delta\gamma_k$ the variation of the risk aversion between $T + (k-1)\,\Delta T$ and $T + (k-1)\,\Delta T$,.and k an integer denoting the number of period.

Proof.

Without loss of generality we assume that k is equal to one. When the number of historical returns T increases of ΔT, the parameter α turns down of $d_T(\alpha)$ where:

$$(I) \qquad d_T(\alpha) = \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{4(\gamma + \Delta\gamma)^2 (T - N - 2)^2} \Delta T \tag{A.4.1}$$

and the sample covariance matrix tends to move closer of the covariance matrix. When the risk aversion γ decreases between time T to time $T+1$ of $\Delta\gamma$, the parameter α increases of $d_\gamma(\alpha)$:

$$(II) \qquad d_\gamma(\alpha) = \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{2\gamma^3(T + \Delta T - N - 2)} \Delta\gamma \tag{A.4.2}$$

and the sample covariance matrix tends to move away of the covariance matrix. If we assume that the increase of α when the risk aversion decreases of $\Delta\gamma$, is higher than the decrease of α when the number of historical returns increases of ΔT, the sample covariance matrix tends to move away from the covariance matrix, we then have:

$$\begin{aligned}
(I) < (II) &\Longrightarrow \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{4(\gamma+\Delta\gamma)^2(T-N-2)^2}\Delta T < \frac{N^2(N+2)(\mu^{\mathbf{T}}\mu)}{2\gamma^3(T+\Delta T-N-2)}\Delta\gamma \\
&\Longrightarrow \frac{\Delta T}{\left[\dfrac{2(\gamma+\Delta\gamma)^2(T-N-2)^2}{\gamma^3(T+\Delta T-N-2)}\right]} < \Delta\gamma \\
&\Longrightarrow \frac{\Delta T}{h} < \Delta\gamma
\end{aligned} \tag{A.4.3}$$

where h is defined as follows:

$$h = \frac{2(\gamma+\Delta\gamma)^2(T-N-2)^2}{\gamma^3(T+\Delta T-N-2)} \tag{A.4.4}$$

we then obtain the result.

■

A.1.5 Proof of Proposition 5.

Proposition 5. When the number of historical returns T goes to infinite, whe have:

$$U(\tilde{\mathbf{\Omega}}) = U(\hat{\mathbf{\Omega}}) \to +\infty$$

Proof.

When T goes to infinite, the utility from our covariance matrix goes to infinite. We also know that the sample covariance matrix is equal to our covariance matrix in this case. As a result we have:

$$U(\hat{\mathbf{\Omega}}_s) = U(\tilde{\mathbf{\Omega}}) \to +\infty \tag{A.5.1}$$

we then obtain the result.

■

A.1.6 Proof of Proposition 6.

Proposition 6. When $T < (N+2)$, the data returns are not enough relevant for the estimation of the covariance matrix, as a result, the naive allocation is the best way for investor. When $(N+2) < T$, the covariance matrix gives to the investor the highest expected utility than the sample covariance matrix for the same level of risk aversion.

Proof.

The expected utility of the covariance matrix is defined as follows:

$$E\left[U(\tilde{\mathbf{\Omega}})\right] = E\left[U(\hat{\mathbf{\Omega}}_s)\right] + \frac{2\gamma(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{A.6.1}$$

which is a sum between the expected utility from the sample covariance matrix and a positive term. We then have:

$$E\left[U(\hat{\mathbf{\Omega}})\right] < E\left[U(\tilde{\mathbf{\Omega}})\right] \tag{A.6.2}$$

We also know that:

$$E\left[U(\hat{\mathbf{\Omega}})\right] = \frac{T}{T-N-2}U(\tilde{\mathbf{\Omega}}) \tag{A.6.3}$$

we then obtain a new expression of $E\left[U(\tilde{\mathbf{\Omega}})\right]$:

$$E\left[U(\tilde{\mathbf{\Omega}})\right] = \frac{T}{T-N-2}U(\tilde{\mathbf{\Omega}}) + \frac{2\gamma}{\mu^{\mathbf{T}}\mu}\frac{T-N-2}{N+2} \tag{A.6.4}$$

Since $(N+2) < T$, we have $U\left(\mathbf{I}\right) < U\left(\tilde{\mathbf{\Omega}}\right)$. We also know that $(T-N-2) < T$, therefore:

$$\begin{aligned}
\left(1 < \frac{T}{T-N-2}\right) &\Longrightarrow \left[U(\mathbf{I}) < \frac{T}{T-N-2}U(\tilde{\mathbf{\Omega}})\right] \\
&\Longrightarrow \left[U(\mathbf{I}) < \frac{T}{T-N-2}U(\tilde{\mathbf{\Omega}}) + \frac{2\gamma}{\mu^{\mathbf{T}}\mu}\frac{T-N-2}{N+2}\right] \\
&\Longrightarrow E\left[U(\mathbf{I})\right] < E\left[U(\tilde{\mathbf{\Omega}})\right]
\end{aligned} \tag{A.6.5}$$

because of the following observation:

$$\begin{cases} 0 < \dfrac{2\gamma}{\mu^{\mathbf{T}}\mu}\dfrac{T-N-2}{N+2} \\ E\left[U(\mathbf{I})\right] = U(\mathbf{I}) \end{cases} \tag{A.6.6}$$

■

A.1.7 Proof of Corollary 1.

Corollary 1. An investor hoping for the tangency portfolio with the maximum ambiguity reliance on the sample tangency portfolio, should not hold this portfolio more than k/h periods, with h defined as follows:

$$h = \frac{4(T-N-2)^2}{T+k-N-2}\sqrt{\frac{T-N-2}{(\mu^{\mathbf{T}}\mu)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}}$$

Proof.

From Proposition 4, by setting $\Delta\gamma$ is equal to zero (what means investor has a constant risk aversion) and ΔT is equall to one, and the risk aversion parameter equals to γ_g where:

$$\gamma_g = \frac{1}{2}\sqrt{\frac{(\mu^{\mathbf{T}}\mu)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{T-N-2}} \tag{A.7.1}$$

we obtain the result.

■

A.1.8 Proof of Corollary 2.

Corollary 2. An investor hoping for the tangency portfolio with a relative ambiguity reliance on the sample tangency portfolio, should not hold this portfolio more than k/h periods, with h defined as follows:

$$h = \frac{4(T-N-2)^2}{(N+2)(T+k-N-2)} \sqrt{\frac{T(T-N-2)}{(\mu^{\mathbf{T}}\mu)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)}}$$

where k is the number of period from time T, for which the risk aversion of investor remains constant.

Proof.

From Corrolary 1, by setting the risk aversion parameter equals to γ_u, where:

$$\gamma_u = \frac{1}{2}\sqrt{\frac{N+2}{T}} \sqrt{\frac{\left(\mu^{\mathbf{T}}\mu\right)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{(T-N-2)}} \tag{A.8.1}$$

we obtain the result.

■

A.1.9 Proof of Proposition 8.

Proposition 8. The expected utility of an investor with a relative ambiguity is higher than the expected utility of an investor having the maximum ambiguity reliance on the sample tangency portfolio.

Proof.

To show this result, we find the risk aversion parameters for which the expected utility of $\tilde{\mathbf{\Omega}}_u$ is higher than the expected utility of any covariance matrix $\tilde{\mathbf{\Omega}}$. Let the following inequality:

$$\left(E\left[U(\tilde{\mathbf{\Omega}}_u)\right] < E\left[U(\tilde{\mathbf{\Omega}})\right]\right) \implies (III) \tag{A.9.1}$$

with:

$$(III) \equiv \frac{T^2 + (N+2)(2T-N-2)}{\sqrt{T(N+2)}}\sqrt{Z} < \frac{T(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}^{-1}\mu)}{2\gamma(T-N-2)} + \frac{2\gamma(2T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{A.9.2}$$

where Z is defined as follows:

$$Z = \frac{\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}^{-1}\mu}{(\mu^{\mathbf{T}}\mu)(T-N-2)(N+2)} \tag{A.9.3}$$

We then obtain:

$$\begin{aligned}(III) &\implies 0 < \gamma^2\left[\frac{2(2T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)}\right] - \gamma\left[\frac{T^2+(N+2)(2T-N-2)}{\sqrt{T(N+2)}}\sqrt{Z}\right] + \left[\frac{T(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}^{-1}\mu)}{2(T-N-2)}\right] \\ &\implies 0 < \mathcal{P}(\gamma)\end{aligned} \tag{A.9.4}$$

where $\mathcal{P}(\gamma)$ denotes an order two polynome defined as follows:

$$\mathcal{P}(\gamma) = A\gamma^2 + B\gamma + C \tag{A.9.5}$$

with:

$$\begin{cases} A = \dfrac{2(2T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \\ B = \dfrac{T^2+(N+2)(2T-N-2)}{\sqrt{T(N+2)}}\sqrt{Z} \\ C = \dfrac{T(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}^{-1}\mu)}{2(T-N-2)} \end{cases} \tag{A.9.6}$$

In order to study the sign of $\mathcal{P}(\gamma)$, we must compute the determinant of the polynome:

$$\begin{aligned} \Delta &= B^2 - 4AC \\ &= \frac{Z\left[T^2+(N+2)(2T-N-2)\right]^2}{T(N+2)} - 4\left[\frac{2(2T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)}\right]\left[\frac{T(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}^{-1}\mu)}{2(T-N-2)}\right] \\ &= \frac{Z\left[T^2+(N+2)(2T-N-2)\right]^2}{T(N+2)} - 4ZT(2T-N-2) \\ &= \frac{Z}{T(N+2)}\left[\left(T^2+(N+2)(2T-N-2)\right)^2 - 4T^2(N+2)(2T-N-2)\right] \\ &= \frac{ZD}{T(N+2)} \end{aligned} \tag{A.9.7}$$

where D is defined as follows:

$$\begin{aligned} D &= \left(T^2+(N+2)(2T-N-2)\right)^2 - 4T^2(N+2)(2T-N-2) \\ &= T^2 + 2T^2(N+2)(2T-N-2) + (2T-N-2) - 4T^2(N+2)(2T-N-2) \\ &= T^2 - 2T^2(N+2)(2T-N-2) + (2T-N-2) \\ &= \left(T^2-(N+2)(2T-N-2)\right)^2 \end{aligned} \tag{A.9.8}$$

We then have:

$$\sqrt{\Delta} = \left(T^2-(N+2)(2T-N-2)\right)\sqrt{\frac{Z}{T(N+2)}} \tag{A.9.9}$$

and we know that $\mathcal{P}(\gamma)$ has two roots γ_1 and γ_2 defined as follows:

$$\begin{cases} \gamma_1 = \dfrac{1}{2}\sqrt{\dfrac{N+2}{T}}\sqrt{\dfrac{\left(\mu^{\mathbf{T}}\mu\right)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{(T-N-2)}} \\ \gamma_2 = \dfrac{1}{2}\dfrac{T^2}{2T-N-2}\sqrt{\dfrac{\left(\mu^{\mathbf{T}}\mu\right)(\mu^{\mathbf{T}}\hat{\mathbf{\Omega}}_s^{-1}\mu)(N+2)}{(T-N-2)}} \end{cases} \tag{A.9.10}$$

We observe that γ_1 is actually equals to the risk aversion parameter corresponding to the covariance matrix with the minimum uncertainty:

$$\begin{cases} \gamma_1 = \sqrt{\dfrac{N+2}{T}}\gamma_g \\ \gamma_2 = \dfrac{T^2}{2T-N-2}\gamma_g \end{cases} \tag{A.9.11}$$

We know that $\mathcal{P}(\gamma)$ has the opposite sign of A inside the roots and the same sign than A outside the roots.

We show now that γ_1 is lower than γ_2. Since $T > 1$, the following inequality is true:

$$\begin{aligned} (T^2 - 2T > 0) &\implies (T^2 - 2T + N + 2 > 0) \\ &\implies 2T - N - 2 < T^2 \\ &\implies 1 < \frac{T^2}{2T-N-2} \\ &\implies 1 < \gamma_2 \\ &\implies \gamma_g < \gamma_2 \end{aligned} \tag{A.9.12}$$

Since $(N+2) < T$, we have $\gamma_1 < 1$, and as a result $\gamma_1 < \gamma_2$.

Since γ_g is the higher risk aversion, because of the decreasing of the first derivative of the utility function $U(\tilde{\mathbf{\Omega}})$ with respect to γ, we can not have a risk aversion parameter higher than γ_g. Since $\gamma_g < \gamma_2$, the risk aversion parameter γ_2 is not relevant. We have the following behavior:

$$\begin{cases} E\left[U(\tilde{\mathbf{\Omega}})\right] < E\left[U(\tilde{\mathbf{\Omega}}_u)\right] & \text{for} \quad \gamma_1 < \gamma \\ E\left[U(\tilde{\mathbf{\Omega}}_u)\right] < E\left[U(\tilde{\mathbf{\Omega}})\right] & \text{for} \quad \gamma < \gamma_1 \end{cases} \tag{A.9.13}$$

Finally, we know that:

$$\gamma_1 = \sqrt{\frac{N+2}{T}}\gamma_g \tag{A.9.14}$$

and since $(N+2) < T$, we have $\gamma_1 < \gamma_g$. As a result, we have:

$$E\left[U(\tilde{\mathbf{\Omega}}_g)\right] < E\left[U(\tilde{\mathbf{\Omega}}_u)\right] \tag{A.9.15}$$

we then obtain the result.

■

A.1.10 Proof of Proposition 9.

Proposition 9. The distance between the tangency portfolio of an investor with the minimum ambiguity and the sample tangency portfolio, is lower than the distance for an investor with a relative ambiguity, which is lower than the distance for an investor with the maximum ambiguity.

Proof.

To show that, we must compare C_g, C_u and C_t with:

$$\begin{cases} C_u = \dfrac{T+N+2}{T} \\ C_g = 2 \\ C_t = \dfrac{2T}{2T-N-2} \end{cases} \tag{A.10.1}$$

Since $(N+2) < T$, we have $C_t < C_g$.

We assume now that $C_u < C_t$, we have the following equation:

$$\begin{aligned} C_u < C_t &\implies \frac{T+N+2}{T} < \frac{2T}{2T-N-2} \\ &\implies \frac{T+N+2}{T} < \frac{2T}{2T-N-2} \\ &\implies (T+N+2)(2T+N+2) < 2T^2 \\ &\implies T < N+2 \end{aligned} \tag{A.10.2}$$

which is wrong. Therefore we have $C_t < C_u$.

■

A.1.11 Proof of Corollary 3.

Corollary 3. An investor hoping for the tangency portfolio with the minimum ambiguity reliance on the sample tangency portfolio, should not hold this portfolio more than k/h periods, with h defined as follows:

$$h = \frac{4(T-N-2)^2}{(N+2)(T+k-N-2)} \sqrt{\frac{(2T-N-2)(T-N-2)}{(\mu^{\mathbf{T}} \hat{\mathbf{\Omega}}_s^{-1} \mu)(\mu^{\mathbf{T}} \mu)}}$$

where k is the number of period from time T, for which the risk aversion of investor remains constant.

Proof.

From Corrolary 1, by setting the risk aversion parameter equals to γ_t, where:

$$\gamma_t = \frac{1}{2} \sqrt{\frac{N+2}{2T-N-2}} \sqrt{\frac{\left(\mu^{\mathbf{T}} \mu\right) (\mu^{\mathbf{T}} \hat{\mathbf{\Omega}}_s^{-1} \mu)(N+2)}{(T-N-2)}} \tag{A.11.1}$$

we obtain the result.

■

A.1.12 Proof of Proposition 10.

Proposition 10. The expected utility of an investor having the minimum ambiguity with the sample mean-variance framework is higher than the expected utility of an investor having a relative ambiguity.

Proof.

We just have to compare the risk aversion for the covariance matrix $\tilde{\mathbf{\Omega}}_t$ with γ_1 because of the following inequality obtained from Proposition 8:

$$\begin{cases} E\left[U(\tilde{\mathbf{\Omega}})\right] < E\left[U(\tilde{\mathbf{\Omega}}_u)\right] & \text{for} \quad \gamma_1 < \gamma \\ E\left[U(\tilde{\mathbf{\Omega}}_u)\right] < E\left[U(\tilde{\mathbf{\Omega}})\right] & \text{for} \quad \gamma < \gamma_1 \end{cases} \tag{A.12.1}$$

Remember, the corresponding risk aversion parameters are defined as follows:

$$\begin{cases} \gamma_1 = \sqrt{\dfrac{N+2}{T}}\gamma_g \\ \gamma_t = \sqrt{\dfrac{N+2}{2T-N-2}}\gamma_g \end{cases} \tag{A.12.2}$$

Since $(N+2) < T$, it is straightforward to see that $T < (2T - N - 2)$. As a result we have:

$$(\gamma_t < \gamma_1) \implies \left(E\left[U(\tilde{\mathbf{\Omega}}_u)\right] < E\left[U(\tilde{\mathbf{\Omega}}_t)\right]\right) \tag{A.12.3}$$

we then obtain the result.

■

A.1.13 Proof of Proposition 11.

Proposition 11. The expected utility of the investor having a relative ambiguity is lower than the expected utility of any investor whose the covariance matrix is denoted by $\tilde{\mathbf{\Omega}}_{gt}^{\beta}$:

$$E\left[U(\tilde{\mathbf{\Omega}}_u)\right] \leq E\left[U(\tilde{\mathbf{\Omega}}_{gt}^{\beta})\right]$$

Proof.

If we assume the opposite of the result we are trying to show, we than have:

$$\begin{aligned} E\left[U(\tilde{\mathbf{\Omega}}_{gt}^{\beta})\right] < E\left[U(\tilde{\mathbf{\Omega}}_u)\right] &\iff \gamma_u < \gamma_{gt}^{\beta} \\ &\iff \gamma_u < \beta\gamma_g + (1-\beta)\gamma_t \\ &\iff \gamma_u - \gamma_t \leq \beta(\gamma_g - \gamma_t) \\ &\iff \frac{\gamma_u - \gamma_t}{\gamma_g - \gamma_t} \leq \beta \\ &\iff \frac{\sqrt{T(N+2)(2T-N-2)} - \sqrt{N+2}}{\sqrt{2T-N-2} - \sqrt{N+2}} \leq \beta \end{aligned} \tag{A.13.1}$$

Since $(N+2) < T$, we know that: $\sqrt{2T-N-2} < T(N+2)(2T-N-2)$, as a result we have:

$$\left(1 < \frac{\sqrt{T(N+2)(2T-N-2)} - \sqrt{N+2}}{\sqrt{2T-N-2} - \sqrt{N+2}} \leq \beta\right) \iff (1 < \beta) \tag{A.13.2}$$

which can not happened, since $0 < \beta < 1$. We then have the following result:

$$E\left[U(\tilde{\mathbf{\Omega}}_u)\right] < E\left[U(\tilde{\mathbf{\Omega}}_{gt}^{\beta})\right] \tag{A.13.3}$$

we then obtain the result.

■

A.1.14 Proof of Proposition 12.

Proposition 12. The utility function of the investor having a relative ambiguity is lower than the utility function of any investor whose the covariance matrix is denoted by $\tilde{\mathbf{\Omega}}_{gt}^{\beta}$:

$$U(\tilde{\mathbf{\Omega}}_u) \leq U(\tilde{\mathbf{\Omega}}_{gt}^{\beta})$$

Proof.

From Proposition 11, the expected utility obtained from any combination between $\tilde{\mathbf{\Omega}}_g$ and $\tilde{\mathbf{\Omega}}_t$ is higher than the one obtained from the covariance matrix $\tilde{\mathbf{\Omega}}_u$:

$$(IV) \equiv \left(E\left[U(\tilde{\mathbf{\Omega}}_u)\right] < E\left[U(\tilde{\mathbf{\Omega}}_{gt}^{\beta})\right]\right) \tag{A.14.1}$$

Remember, the expression for the expected utility depends on the expression for the utility:

$$E\left[U(\tilde{\mathbf{\Omega}})\right] = \frac{T}{T-N-2}U(\tilde{\mathbf{\Omega}}) + \frac{2\gamma(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} \tag{A.14.2}$$

We then have from $(A.14.1)$:

$$\begin{aligned}(IV) \iff & \left[\frac{T}{T-N-2}U(\tilde{\mathbf{\Omega}}_u) + \frac{2\gamma_u(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)} < \frac{T}{T-N-2}U(\tilde{\mathbf{\Omega}}_{gt}^{\beta}) + \frac{2\gamma_{gt}^{\beta}(T-N-2)}{(\mu^{\mathbf{T}}\mu)(N+2)}\right] \\ \iff & U(\tilde{\mathbf{\Omega}}_u) - U(\tilde{\mathbf{\Omega}}_{gt}^{\beta}) < \frac{2(T-N-2)^2}{T(\mu^{\mathbf{T}}\mu)(N+2)}\left(\gamma_{gt}^{\beta} - \gamma_u\right)\end{aligned} \tag{A.14.2}$$

Since (IV) is true, we also have $\gamma_{gt}^{\beta} < \gamma_u$, what means $\gamma_{gt}^{\beta} - \gamma_u < 0$. As a result:

$$\left[U(\tilde{\mathbf{\Omega}}_u) - U(\tilde{\mathbf{\Omega}}_{gt}^{\beta}) < \frac{2(T-N-2)^2}{T(\mu^{\mathbf{T}}\mu)(N+2)}\left(\gamma_{gt}^{\beta} - \gamma_u\right) < 0\right] \iff U(\tilde{\mathbf{\Omega}}_u) < U(\tilde{\mathbf{\Omega}}_{gt}^{\beta}) \tag{A.14.2}$$

we then obtain the result.

■

A.1.15 Proof of Corollary 4.

Corollary 4. The optimal parameter β^* for the linear combination between investors having the maximum ambiguity and investors having the minimum ambiguity reliance on the Sahrpe model is defined as follows:

$$\beta^* = \sqrt{\frac{N+2}{T}} \frac{\sqrt{2T-N-2} - \sqrt{T}}{\sqrt{2T-N-2} - \sqrt{N+2}}$$

Proof.

Since $\tilde{\mathbf{\Omega}}_u$ is the optimal combination between $\tilde{\mathbf{\Omega}}_g$ and $\tilde{\mathbf{\Omega}}_t$, the following equality allows for finding the optimal β:

$$\begin{aligned}
\left[\tilde{\mathbf{\Omega}}_{gt}^{\beta} = \tilde{\mathbf{\Omega}}_u\right] &\iff \left[\beta^*\tilde{\mathbf{\Omega}}_g + (1-\beta^*)\tilde{\mathbf{\Omega}}_t = \tilde{\mathbf{\Omega}}_u\right] \\
&\iff \beta^*\gamma_g + (1-\beta^*)\gamma_t = \gamma_u \\
&\iff \beta^*\gamma_g + (1-\beta^*)\gamma_g\sqrt{\frac{N+2}{2T-N-2}} = \sqrt{\frac{N+2}{T}}\gamma_g \\
&\iff \beta^*\left(1 - \sqrt{\frac{N+2}{2T-N-2}}\right) = \sqrt{\frac{N+2}{T}} - \sqrt{\frac{N+2}{2T-N-2}} \\
&\iff \beta^*\left(\frac{\sqrt{2T-N-2} - \sqrt{N+2}}{\sqrt{2T-N-2}}\right) = \sqrt{N+2}\left(\frac{\sqrt{2T-N-2} - \sqrt{T}}{\sqrt{T(2T-N-2)}}\right) \\
&\iff \beta^* = \sqrt{\frac{N+2}{T}}\left(\frac{\sqrt{2T-N-2} - \sqrt{T}}{\sqrt{2T-N-2} - \sqrt{N+2}}\right)
\end{aligned} \tag{A.15.1}$$

we then obtain the result.

■

A.2 List of Figures

Figure 1: The Shape of the Optimal Parameter

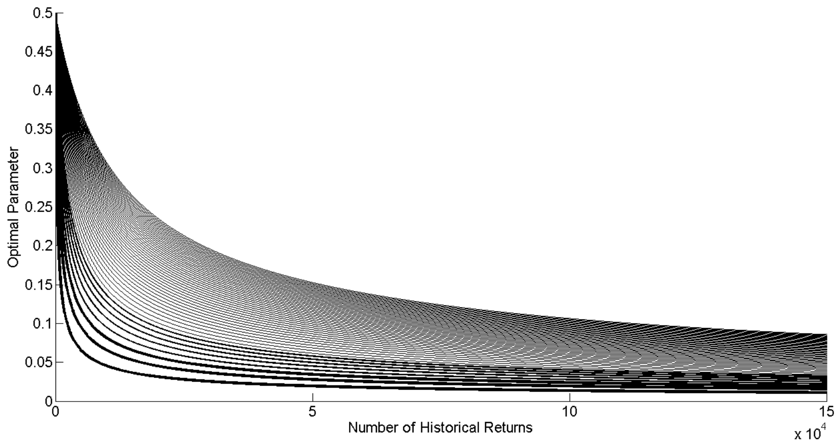

Figure 1: **Source**: *Evolution of the optimal parameter with respect to the number of historical returns, bolder the line and higher the number of assets, the first window size is obtained by adding three to the number of assets, computation by the authors.*

Figure 2: Illustration of the Power Law Character of the Ambiguity

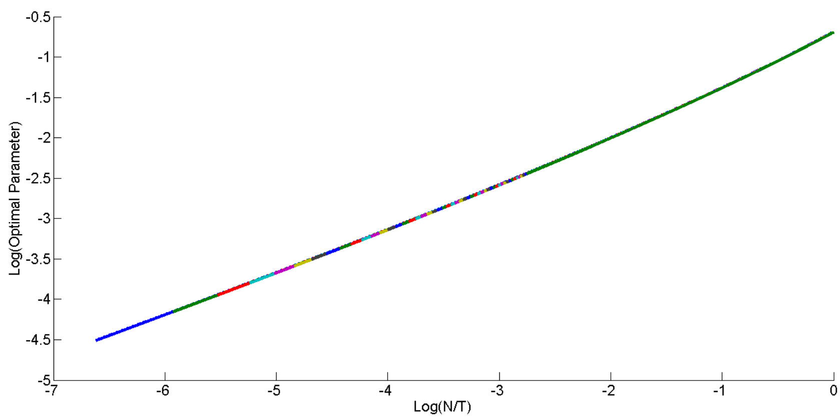

Figure 2: **Source**: *Evolution of the optimal parameter with respect to the size of the invesment universe in logarithm, the colours denote the shape for a given number of assets, the first window size is obtained by adding three to the number of assets, the slope of the straight lines is around equals to 0.58, computation by the authors.*

Figure 3: The Dynamic of the Risk Aversion Parameter

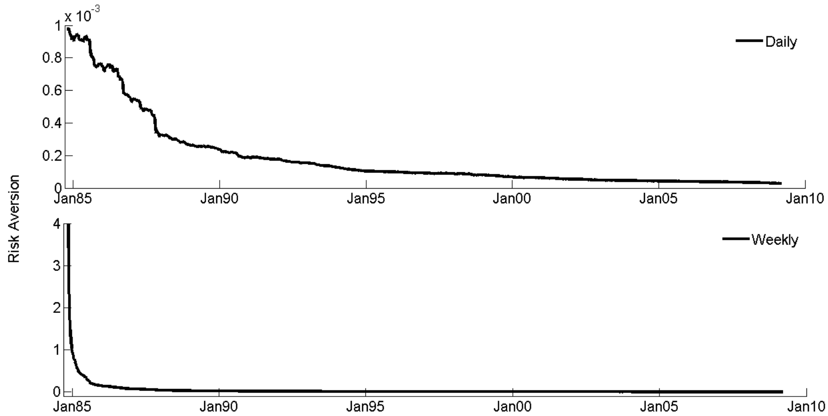

Figure 3: **Source**: *Dynamic of the risk aversion parameter with respect to the size of the invesment universe, 250 assets from the S&P500 universe, 253 historical returns for the first estimation window when considering the weekly frequency, the number of historical returns for the daily frequency is obtained by considering the same window date as for the weekly frequency, the sample window increases of one frequency until the end date, computation by the authors.*

Figure 4: Comparison Between the Volatility of the Risk Aversion Parameter and the S&P500 Index in a Daily Frequency

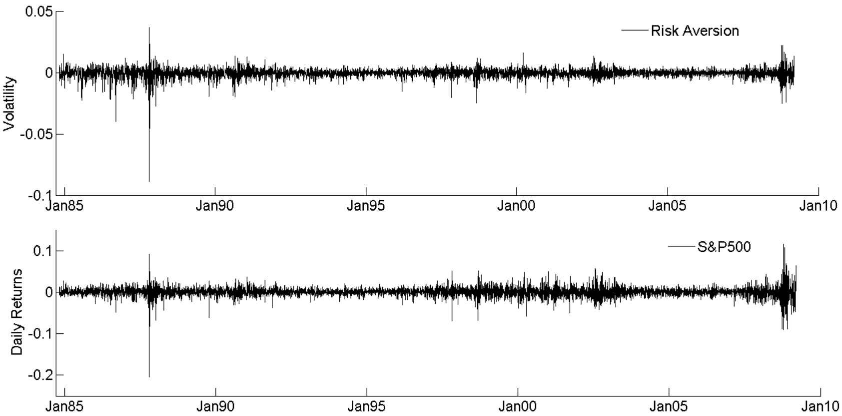

Figure 4: **Source**: *Volatility of the risk aversion parameter and volatility of the S&P500 index from 01/1985 to 03/2009, 250 assets from the S&P500 universe, daily frequency, computation by the authors.*

Figure 5: Comparison Between the Volatility of the Risk Aversion Parameter and the S&P500 Index in a Weekly Frequency

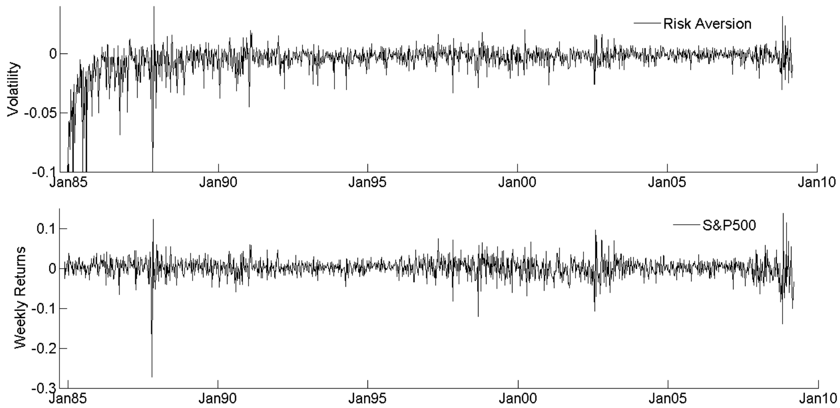

Figure 5: **Source**: *Volatility of the risk aversion parameter and volatility of the S&P500 index from 01/1985 to 03/2009, 250 assets from the S&P500 universe, weekly frequency, computation by the authors.*

Figure 6: The Contribution of the Sample Size Neglect to the Market Volatility

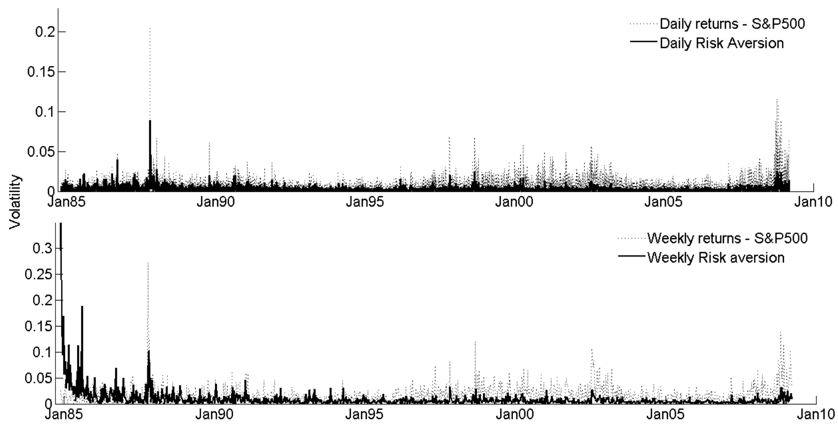

Figure 6: **Source**: *Volatility of the risk aversion parameter and volatility of the S&P500 index from 01/1985 to 03/2009, 250 assets from the S&P500 universe, the daily frequency on the top, the weekly frequency bottom, computation by the authors.*

Figure 7: Sharpe Ratio of Tangency Portfolios

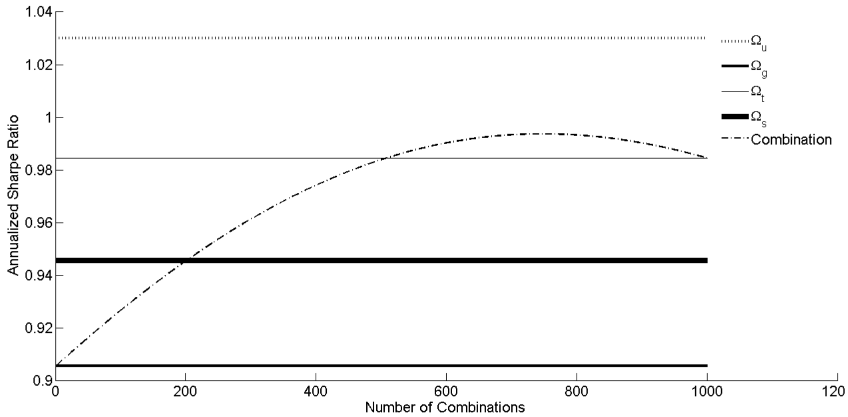

Figure 7: **Source**: *Datastream, Sharpe ratios of the out-of-sample tangency portfolios, the sample window increases from 326 to 1004 historical returns, 717 periods of estimation window, 324 assets of the S&P500 universe, no completion need, weekly frequency, computation by the authors.*