Optimal Portfolio Rule:
When There is Uncertainty in The
Parameter Estimates

by

Hyunjong Jin

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Quantitative Finance

Waterloo, Ontario, Canada, 2012

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The classical mean-variance model, proposed by Harry Markowitz in 1952, has been one of the most powerful tools in the field of portfolio optimization. In this model, parameters are estimated by their sample counterparts. However, this leads to estimation risk, which the model completely ignores. In addition, the mean-variance model fails to incorporate behavioral aspects of investment decisions. To remedy the problem, the notion of ambiguity aversion has been addressed by several papers where investors acknowledge uncertainty in the estimation of mean returns. We extend the idea to the variances and correlation coefficient of the portfolio, and study their impact. The performance of the portfolio is measured in terms of its Sharpe ratio. We consider different cases where one parameter is assumed to be perfectly estimated by the sample counterpart whereas the other parameters introduce ambiguity, and vice versa, and investigate which parameter has what impact on the performance of the portfolio.

Acknowledgements

Contents

Chapter 1

Introduction

The mean-variance (MV) model, first introduced by Harry Markowitz [Markowitz (1952)], has been popular since its inception, and the main concept of the model has been widely used by mutual and pension funds. The model requires knowledge of the expected returns, variances, and covariances of the returns of the assets in the portfolio. However, in practice, these parameters are typically estimated from a sample, and the estimation risk is ignored. Despite the usefulness of the MV model and its simplicity, the presence of estimation risk is a major limitation, and many researchers have noted its effects on the choice of portfolio rules. The estimates are solely based on historical performance, and Fabozzi stresses that "markets and economic conditions change throughout time", so past performance is a poor indication of future returns, [Fabozzi, (2007)]. In addition, Michaud suggests that "MV optimization significantly overweights (underweights) those securities that have large (small) estimated returns, negative (positive) correlations and small (large) variances", [Michaud (1989)]. Kroll and Levy study the effect of sampling error on the portfolio, and conclude that a large sample size is required to maintain the extent of the error at a prescribed level, [Kroll, Levy (1980)].

Another practical limitation of the classical mean-variance model is "the computational difficulty associated with solving large-scale quadratic programming problem with a dense covariance matrix", [Konno, Yamazaki (1991)]. Konno and Yamazaki propose the mean-

absolute deviation (MAD) portfolio optimization model to circumvent in the difficulty the pain of dealing with quadratic problems. The mean-absolute deviation model retains all the features of the mean-variance model, and does not require the covariance matrix. However, Simann shows that ignoring the covariance matrix leads to greater estimation risk, [Simann (1997)].

An effort to control estimation risk leads to the Bayesian approach, which is more practical. The discrepancy between population and sample parameters is acknowledged by treating unknown parameters as random variables. A Bayesian investor constructs a predictive distribution from a pre-specified distribution with observations, but is assumed to be neutral with respect to Knightian uncertainty, [Knight (1921)]. In this context, risk refers to measurable risk that can be represented numerically, whereas uncertainty refers to unmeasurable risk which cannot. There is extensive literature showing that a rational investor does not treat risk and uncertainty the same way, and that aversion to such uncertainty is present in decision making. This is best illustrated by the Ellsberg paradox, [Ellsberg (1961)]. The Bayesian approach fails to capture this aversion, which is commonly referred to as ambiguity aversion, [Epstein (1999)].

Uncertain of returns certaintly cannot be overlooked because it "tends to have more influence than risk in mean-variance optimiation", [Fabozzi (2007)]. Thus, the need to consider investors with multiple priors arises to take ambiguity aversion into consideration, and Garlappi, Uppal, and Wang propose a modified mean-variance model which is based on a max-min optimization of the utility function, [Garlappi, Uppal, Wang (2007)]. The usage of confidence intervals is a widely used concept in robust optimization, and it requires the knowledge of the underlying distribution of the data, [Fabozzi (2007)]. The model proposed by Garlappi et. al allows uncertainty in mean returns by obtaining confidence intervals instead of point estimates. However, they still assume that the covariance matrix of the assets is perfectly estimated by its sample counterpart. Just like the estimate for sample mean returns, the sample covariance matrix is prone to estimation risk, which many researchers have tried to improve by applying different techniques, such as shrinkage, to covariance estimation. Fabozzi also suggests robust estimation of covariance matrices using elliptic distributions.

In this thesis, we present a model that incorporates uncertainty in the covariance matrix as well as mean returns of the assets, and study the impact of this new layer of uncertainty, and aim

1. to explain why ambiguity aversion could be useful in the mean-variance model,

2. to study the impact of ambiguity aversion in other model parameters,

3. to investigate which parameter has the greatest impact on the portfolio performance.

Chapter 2 re-visits the standard mean-variance model, and further discusses the behavioral aspects of portfolio optimization. In Chapter 3, we motivate our problem by considering the simplest case where the portfolio consists of one risky asset and one risk-free asset. In this simple case, we analyze how uncertainty in mean returns and the variances can affect the portfolio rules. In Chapter 4, we extend the simplest case presented in Chapter 3 to the multi-asset case. A complication arises in the multi-asset case with the presence of correlations among the risky assets. We study the distribution of correlation coefficients, and analyze the portfolio rules. Chapter 5 applies theoretical results from earlier sections to empirical data to verify our model and analysis. We conclude with recommendations on how to extend the idea and results to more general cases.

Chapter 2

Motivation

Mean-variance analysis is the cornerstone of modern portfolio theory. Although the extensive literature claims that the mean-variance model itself is not practical, its core concept has been widely used, and it is worth reviewing the model. In the first section, we revisit the classical mean-variance portfolio model, and discuss mathematically how sample estimates of the parameters in the model may lead to poor results/performance of the portfolio.

The biggest drawback in the classical mean-variance model is the lack of the behavioral aspects of decision making. In practice, information available to investors is imperfect and investors are prone not only to risk, but also to uncertainty in their choice of portfolio rules. Existing research mostly focuses on dealing with risk, but there is evidence that investors react to uncertainty in ways that violate the expected utility hypothesis. In the second section, we first study the expected utility hypothesis, and present the Ellesberg paradox. This famous example violates the expected utility hypothesis, and we discuss its implications on portfolio rules.

2.1 Mean-Variance Optimal Allocation

The classical mean-variance model suggests that the optimal portfolio rule of N stocks (or any risky assets) is obtained by solving the optimization problem

$$\max_{\omega} \left[\omega^T \mu - \frac{\gamma}{2} \omega^T \Sigma \omega \right] \tag{2.1}$$

where $\omega \in \mathbb{R}^N$ denotes proportions of wealth invested in each stock, $\gamma > 0$ is a constant that represents the investors' risk aversion, $\mu \in \mathbb{R}$ is the true excess returns over the risk-free asset, and $\Sigma \in \mathbb{R}^{N \times N}$ is the covariance matrix of the N stocks, which is positive definite.

Proposition 2.1.1 *The solution to the optimization problem* (2.1) *is given by*

$$\omega^* = \frac{1}{\gamma} \Sigma^{-1} \mu \tag{2.2}$$

where ω^* denotes the optimal portfoio rule. When there is no risk-free asset, the solution to (2.1) *is*

$$\omega^{**} = \frac{1}{\gamma} \Sigma^{-1} \left(\mu - \mu_0 \cdot \mathbb{1} \right) \tag{2.3}$$

where

$$\mu_0 = \frac{\mu^T \Sigma^{-1} \mathbb{1} - \gamma}{\mathbb{1}^T \Sigma^{-1} \mathbb{1}} \tag{2.4}$$

is the expected return on the zero-beta portfolio associated with the optimal portfolio ω, and $\mathbb{1} \in \mathbb{R}^N$ is a N-vector of 1's.

Proposition 2.1.1 is a well-known result, but we recall here the proof of it.

Proof. With no risk-free asset, the additional constraint $\omega^T \mathbb{1} = 1$ is imposed. Let $\lambda > 0$. By the method of Lagrange multipliers, we define

$$\Lambda (\omega, \lambda) = \omega^T \mu - \frac{\gamma}{2} \omega^T \Sigma \omega - \lambda \left(\omega^T \mathbb{1} - 1 \right)$$

First order conditions are

$$\frac{\partial \Lambda(\omega, \lambda)}{\partial \omega} = \mu - \gamma \Sigma \omega - \lambda \mathbb{1}$$

$$\frac{\partial \Lambda(\omega, \lambda)}{\partial \lambda} = -\omega^T \mathbb{1} + 1$$

By setting the first order conditions to 0, we obtain

$$\omega = \frac{1}{\gamma} \Sigma^{-1} (\mu - \lambda \mathbb{1})$$

$$1 = \omega^T \mathbb{1}$$

We note that $\omega^T \mathbb{1} = \mathbb{1}^T \omega$, so

$$1 = \mathbb{1}^T \left[\frac{1}{\gamma} \Sigma^{-1} (\mu - \lambda \mathbb{1}) \right]$$

$$\implies \gamma = \mathbb{1}^T \Sigma^{-1} \mu - \mathbb{1}^T \Sigma^{-1} \lambda \mathbb{1}$$

$$\implies \lambda = \frac{\mathbb{1}^T \Sigma^{-1} \mu - \gamma}{\mathbb{1}^T \Sigma^{-1} \mathbb{1}}$$

$$= \frac{\mu^T \Sigma^{-1} \mathbb{1} - \gamma}{\mathbb{1}^T \Sigma^{-1} \mathbb{1}}$$

Choose $\mu_0 = \lambda$ to obtain (2.4), and we obtain the desired result. □

The underlying assumption of the classical mean-variance portfolio model is that investors have perfect information about the market, or in other words, μ and Σ are known with certainty. Since this is not true in practice, investors estimate the parameters μ and Σ by their counterparts, $\hat{\mu}$, the sample mean, and $\hat{\Sigma}$, the sample covariance matrix. Hence, the solution to the classical mean-variance model becomes

$$\hat{\omega} = \frac{1}{\gamma} \hat{\Sigma}^{-1} \hat{\mu} \tag{2.5}$$

If we use T observations for each stock, and if we assume that the excess returns $\mu_i \in \mathbb{R}^N, i = 1, 2, \cdots, T$ are normally distributed, then $\hat{\mu} \sim N\left(\mu, \frac{\Sigma}{T}\right)$, a multinormal distribution with mean $\mu \in \mathbb{R}^N$ and covariance matrix $\Sigma \in \mathbb{R}^{N \times N}$. $(T-1)\hat{\Sigma} \sim W_N(T-1, \Sigma)$, is a Wishart distribution with covariance matrix $\Sigma \in \mathbb{R}^{N \times N}$ and $T-1$ degrees of freedom,

where

$$\hat{\mu} = \frac{1}{T} \sum_{i=1}^{T} \mu_i$$

$$\hat{\Sigma} = \frac{1}{T-1} \sum_{i=1}^{T} (\mu_i - \hat{\mu}) (\mu_i - \hat{\mu})^T$$

In the expression for the sample covariance matrix, we notice that the denominator is $T-1$ instead of T. This is known as Bessel's correction, and is needed for $\hat{\Sigma}$ to be an unbiased estimator of Σ. As $T \to \infty$, the difference between $T-1$ and T becomes negligible, and the significance of Bessel's correction diminishes.

Proposition 2.1.2 *If the sample covariance follows the Wishart distribution with $T-1$ degrees of freedom and covariance matrix Σ, or mathematically, $\hat{\Sigma} \sim W_N (T-1, \Sigma)$, then*

$$\mathbb{E}\left[\hat{\Sigma}^{-1}\right] = \frac{\Sigma^{-1}}{T-N-2}$$

The result follows from the fact that $\hat{\Sigma}^{-1} \sim W_N^{-1} (T-1, \Sigma^{-1})$, the inverted Wishart distribution. One fundamental property of a normal distribution is that the sample mean and sample covariance are independent. In other words,

$$\mathbb{E}\left[\hat{\Sigma} \cdot \hat{\mu}\right] = \mathbb{E}\left[\hat{\Sigma}\right] \cdot \mathbb{E}[\hat{\mu}]$$

If we take the expectation of $\hat{\omega}$ in (2.5),

$$\mathbb{E}[\hat{\omega}] = \frac{T-1}{T-N-2}\omega^* \tag{2.6}$$

where ω^* is the optimal portfolio rule with perfection information in (2.2).

Proposition 2.1.2 is a standard statistics result that can be found in [Morrison (1967)]. In Proposition 2.1.2, $\hat{\omega}$ is a vector, and we define $\mathbb{E}[\hat{\omega}]$ as the vector of the expectation of each element of $\hat{\omega}$. Provided that $T > N - 2$, we obtain $|\mathbb{E}[\hat{\omega}]| > |\omega^*|$ from (2.6), where the inequality sign is component-wise. This result indicates that investors who use the sample mean and sample covariance as true parameters tend to overestimate the true optimal portfolio rule, which may lead to poor performance out of sample.

7

2.2   Behavioral Aspects of Portfolio Choice

The expected utility hypothesis was first initiated by Bernoulli in 1738, and incorporates risk aversion in the preferences of people with regard to unpredictable outcomes. Daniel Bernoulli stated

> "The determination of the value of an item must not be based
> on the price, but rather on the utility it yields. There is
> no doubt that a gain of one thousand ducats is more
> significant to the pauper than to a rich man though both gain
> the same amount."

The expected value criterion is a naive rule that prefers an investment with higher expected value. However, the St.Petersburg paradox shows a contradiction between the choice that rational investors make and the choice that the expected value criterion recommends. We consider a simple example to illustrate this point.

Example 2.2.1 *A fair coin is tossed repeatedly until a tail appears, which marks the end of the game. If a tail appears on k^{th} toss for the first time, you win 2^{k-1}. For example, if a head appears on the first toss and then a tail on the second toss, you win $2. How much would you pay to enter this game?*

If we consider the expected payoff of the game presented in Example 2.2.1, we see that it is infinity. In other words, based on the expected value criterion, one should enter the game for any fixed amount of dollars, which is, however, not the case with rational investors. The expected value criterion fails to account for a highly unlikely event with a large payoff that rational (or risk-averse) investors tend to opt out of. The expected utility theory was proposed to remedy this problem of the expected value criterion possesses. The classical mean-variance model is consistent with the expected utility theory since

- The model introduces a constant that represents investors' risk-aversion, denoted as γ in (2.1).

- The utility function that investors aim to maximize is a completely concave function with respect to proportions of wealth allocated in the portfolio.

However, the expected utility theory has been challenged for years with empirical results that violate it. The most famous example is the Ellsberg paradox [Ellsberg (1961)], which we present here to illustrate violations of the expected utility theory.

Example 2.2.2 *You have an urn that contains 30 red balls and 60 other balls that are either black or yellow in unknown proportion. You draw one ball from the urn of a total of 90 balls. Suppose you are given the following two choices to bet on.*

	Red	Black	Yellow
I	$100	$0	$0
II	$0	$100	$0

Choice I means that you bet on 'Drawing a red ball', and Choice II means that you bet on 'Drawing a black ball', with the exact same payout. Which choice would you take? Consider the following two other choices.

	Red	Black	Yellow
III	$100	$0	$100
IV	$0	$100	$100

Choice III means that you bet on 'Drawing either a red or a yellow ball', and Choice IV means that you bet on 'Drawing either a black or a yellow ball'. Which choice would you take?

We denote $a \succ b$ to mean 'a is preferred to b'. A natural response, based on [Ellsberg (1961)], is Choice I $\succ$ Choice II, and Choice IV $\succ$ Choice III. In this simple model (or game), uncertainty comes from unknown proportion between black and yellow balls. The probability of drawing a red ball is $P(\mathrm{I}) = \frac{1}{3}$, and the probability of drawing a black ball is

$P(\text{II}) \in \left[0, \frac{2}{3}\right]$, and the distribution of the number of black balls is even unknown. Choice I $\succ$ Choice II indicates that people prefer a fixed probability to a uncertain probability even though the uncertain one could potentially be higher than the fixed one. In other words, people are averse to such uncertainty or ambiguity. Based on the expected utility theory, Choice I $\succ$ Choice II should imply Choice III $\succ$ Choice IV because the only difference made to the first set of choices (I and II) from the second set of choices (III and IV) is the addition of extra payoff to drawing a yellow ball. Mathematically, $a \succ b \implies a + c \succ b + c, \forall c > 0$. However, the second set of choices clearly violates this property as Choice IV $\succ$ Choice III. This violation happens due to a removal of ambiguity that was present with the first set of choices. The probability of drawing either a black or yellow ball is $P(\text{IV}) = \frac{2}{3}$ with certainty. The second set of choices also possesses ambiguity in Choice III as the probability of drawing either a red or yellow ball is $P(\text{III}) \in \left[\frac{1}{3}, 1\right]$.

Example 2.2.2 shows that there is some other factor that the expected utility theory missed (because clearly, it is violated), and that people are averse to ambiguity. The notion of ambiguity aversion should be distinguished from risk aversion. In simple terms, risk aversion is due to pessimism about events that are less likely to happen whereas ambiguity aversion is associated with situations where one has less knowledge of what he is getting himself into. To elaborate and make a connection with portfolio rules, we consider the mean-variance model. Investors estimate the parameters, and obtain what they think are true values. Based on the variance (risk) of the portfolio, it may or may not perform better than the estimated return. Risk averse investors do not appreciate high variance (risk) and tend to favour porfolios with lower variance. Unfortunately, estimation error is always present, and they cannot be certain how accurate their estimates are. Consider two hypothetical portfolios with their estimated returns and variances in the following table.

	Return	Variance
Portfolio I	10%	30%
Portfolio II	15%	30%

Risk averse investors would definitely choose Portfolio II over Portfolio I. However, investors cannot be 100% certain of these point estimates. Suppose they believe that Portfolio I would have a rate of return between 8% and 20% with variance between 25%

and 35%, and that Portfolio II would have a rate of return between 7% to 18% with variance 25% and 60%. In other words, the 'actual' (true) returns and variances of the two portfolios are uncertain (ambiguous). In this case, such ambiguity would lead risk averse investors to believe their choice of Portfolio II could be prone to more risk for a given rate of return. In the next chapter, we discuss how to treat ambiguity with statistical distributions, and construct a model that incorporates ambiguity aversion.

Chapter 3

Simple World: One-Stock Case

In this thesis, expected returns of risky assets are assumed to follow a normal distribution. At first, the normality assumption on equity returns appears to be restrictive. However, we can make our case by observing that intraday returns consist of many trades over short time periods. Returns over these short periods might not be normally distributed. However, as the number of trades per day is considered to be large enough, based on Central Limit Theorem[1], we suggest that daily returns are approximately normal. Based on this assumption, our model requires knowledge of the normal, χ^2, and Student's t distributions, and we recall these distributions and their associated properties. The main focus of this chapter is the study of a simple portfolio consisting of only one risky asset and a risk-free asset. We also illustrate the impact of uncertainty on portfolio decision rules.

---

[1] The distribution of the average of many random numbers is normally distributed, independent of the distribution of each number.

3.1 Review of Statistics and Distributions

3.1.1 Normal and Student's t Distributions

The probability density function of the normal distribution $N\left(\mu, \sigma^2\right)$ is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R} \tag{3.1}$$

where μ is the mean, and σ^2 is the variance. In particular, the standard normal distribution refers to $N\left(0, 1\right)$, and any normal random variable can be standardized to follow the standard normal distribution. Let $X \sim N\left(\mu, \sigma^2\right)$. Then,

$$Z = \frac{X - \mu}{\sigma} \tag{3.2}$$

is a standard normal random variable.

One useful property of the normal distribution is that the sum of identically independently distributed (i.i.d.) normal random variables also follows a normal distribution. Suppose $X_1, X_2, \cdots, X_n \sim N\left(\mu, \sigma^2\right)$. Then,

$$\sum_{i=1}^{n} X_i \sim N\left(n\mu, n\sigma^2\right) \tag{3.3}$$

and from (3.3), we can easily see that

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \tag{3.4}$$

Another useful property is that the sample mean and the sample (co-)variance are independent, as noted earlier in Proposition 2.1.2. This particular property can be proven with Basu's theorem which states that any bounded, complete sufficient statistic is independent of any ancillary statistic[2]. Suppose $X_1, X_2, \cdots, X_n \sim N\left(\mu, \sigma^2\right)$, i.i.d. normal random

---

[2]An ancillary statistic is a statistic whose sampling distribution does not depend on the (unknown) parameter being sampled. For example, if $X_1, \cdots, X_n$ are i.i.d. normal random variables with unknown expected value μ and known variance σ^2, the sample variance $\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$ is an ancillary statistic because its sampling distribution does not chance as μ changes.

variables. Then, the sample mean

$$\hat{\mu} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{3.5}$$

is a complete sufficient statistic, and the sample variance

$$s^2 = \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2}{n-1} \tag{3.6}$$

is an ancillary statistic, and hence, they are independent.

The Student's t-distribution (or simply t-distribution) is closely related to the normal distribution as the t-distribution arises in estimating the mean of a normally distributed population when the sample size is relatively small. In mathematical terms, we consider a population of i.i.d. normal random variables $\{X_i\}_{i \in \mathbb{N}}$, $X_i \sim N\left(\mu, \sigma^2\right), \forall i \in \mathbb{N}$, and a subset of n elements $\{X_{t_1}, X_{t_2}, \cdots, X_{t_n}\}, t_j \in \mathbb{N}, j = 1, 2, \cdots, n$. Then, the sample mean of the subset, $\hat{\mu}$, follows the t-distribution with $n-1$ degrees of freedom.

The probability density function of the t-distribution is given by

$$f(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi} \cdot \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{1}{2}(\nu+1)} \tag{3.7}$$

where $\nu = n-1$ is degrees of freedom, and $\Gamma(\cdot)$ is the gamma function defined by

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

As the sample size n grows, it can be verified that the t-distribution approaches the standard normal distribution.

3.1.2 χ^2 Distribution

The χ^2 distribution with ν degrees of freedom, often denoted as $\chi^2(\nu)$, is the distribution of a sum of the squares of ν independent standard normal random variables. The probability density function of $\chi^2(\nu)$ is given by

$$f(x) = \frac{1}{2^{\frac{\nu}{2}} \cdot \Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}, \quad x \in \mathbb{R}^+ \tag{3.8}$$

where $\mathbb{R}^+$ denotes a set of all non-negative real numbers.

Remark 3.1.1 *If $X \sim N(0,1)$ and $Y \sim \chi^2(n)$, then*

$$\frac{X}{\sqrt{Y/n}} \sim t_{n-1} \tag{3.9}$$

where t_{n-1} denotes the t-distribution with $n-1$ degrees of freedom. Moreover, X and Y are independent.

We recall another useful property of the χ^2 distribution.

Remark 3.1.2 *Consider a population of $\{X_i\}_{i \in \mathbb{N}}$ $X_i \sim N(\mu, \sigma^2), \forall i \in \mathbb{N}$, and a subset of n elements $\{X_{t_1}, X_{t_2}, \cdots X_{t_n}\}, t_j \in \mathbb{N}, j = 1, 2, \cdots, n$. Then,*

$$(n-1)\frac{s^2}{\sigma^2} \sim \chi^2(n-1) \tag{3.10}$$

where the sample variance s^2 is defined in (3.6).

Remarks 3.1.1 and 3.1.2 are useful building blocks on our model as illustrated in the next section.

3.2   Uncertainty in One-Stock Case

As a simple illustration, we consider a portfolio that consists of one stock (or any other risky asset) and a risk-free asset. Then, (2.1) is simplified as

$$\max_{\omega_1} \left[\omega_1 \mu_1 - \frac{\gamma}{2}\omega_1^2 \sigma_1^2 \right] \tag{3.11}$$

where ω_1 is the proportion of wealth invested in the stock, μ_1 is the true return on the stock, and σ_1^2 is the variance of the stock return. A fundamental idea is to restrict the

return and the variance of the stock to fall within particular intervals, rather than to obtain point estimates. If T denotes the number of observations, then from Remarks 3.1.1 and 3.1.2 with (3.2), we can easily see that

$$\frac{\mu_1 - \hat{\mu}_1}{s_1/\sqrt{T}} \sim t_{T-1} \tag{3.12}$$

and

$$(T-1)\frac{s_1^2}{\sigma_1^2} \sim \chi^2\,(T-1) \tag{3.13}$$

where $\hat{\mu}_1$ and s_1^2 are the sample mean and the sample variance, respectively, on the risky asset. The objective is to control unknown quantities μ_1 and σ_1^2, and (3.12) and (3.13) suggest that we can construct intervals in which those quantities lie, at a prescribed confidence level. We introduce ambiguity parameters $\varepsilon_1 \in [0, \infty)$, $\delta_1^l \in [0,1]$, and $\delta_1^u \in [0,\infty)$ such that

$$\left(\frac{\mu_1 - \hat{\mu}_1}{s_1/\sqrt{T}}\right)^2 \leq \varepsilon_1 \tag{3.14}$$

and

$$1 - \delta_1^l \leq \frac{s_1^2}{\sigma_1^2} \leq 1 + \delta_1^u \tag{3.15}$$

If investors are absolutely certain that perfect information on the market is available to them, or in other words, that returns and variances of the risky assets are estimated by their sample counterparts with no estimation risk, then the ambiguity parameters are equivalent to 0. Non-zero ambiguity parameters reflect an acknowledgement of possible discrepancies between μ_1 and $\hat{\mu}_1$, and between σ_1 and s_1. This can be seen from (3.14) and (3.15) as $\mu_1 = \hat{\mu}_1$ and $\sigma_1 = s_1$ if and only if $\varepsilon_1 = 0, \delta_1^l = 0$, and $\delta_1^u = 0$. In practice, this ideal situation is hardly the case. At a specific confidence level α, we can obtain the values of ε_1, δ_1^l, and δ_1^u as the statistics in (3.14) and (3.15) follow well-known distributions whose cumulative distribution functions are readily available.

Once we construct confidence intervals for μ_1 and σ_1, we impose an additional minimization over the set of possible values for μ_1 and σ_1, subject to (3.14) and (3.15). This

additional minimization ensures that investors are not neutral to ambiguity, [Garlappi, Uppal, Wang (2007)]. In other words, it reflects the behavior of investors that they tend to invest less (more) as the ambiguity level rises (lowers).

The problem (3.11) can be formulated as

$$\max_{\omega_1} \left[\min_{\mu_1, \sigma_1} \left[\omega_1 \mu_1 - \frac{\gamma}{2} \omega_1^2 \sigma_1^2 \right] \right] \tag{3.16}$$

subject to

$$\begin{aligned} \hat{\mu}_1 - \frac{s_1 \sqrt{\varepsilon_1}}{\sqrt{T}} &\leq \mu_1 \leq \hat{\mu}_1 + \frac{s_1 \sqrt{\varepsilon_1}}{\sqrt{T}} \\ \frac{s_1^2}{1 + \delta_1^u} &\leq \sigma_1^2 \leq \frac{s_1^2}{1 - \delta_1^l} \end{aligned} \tag{3.17}$$

Remark 3.2.1 *We note that* $\text{sign}(\hat{\mu}_1)$ *is known since* $\hat{\mu}_1$ *is computed directly from historical observations. From the constraints* (3.17), *we see that the necessary condition to determine* $\text{sign}(\mu)$ *is*

$$|\hat{\mu}_1| > \frac{s_1 \sqrt{\varepsilon_1}}{\sqrt{T}} \tag{3.18}$$

Otherwise, $\hat{\mu}_1 - \frac{s_1 \sqrt{\varepsilon_1}}{\sqrt{T}}$ *and* $\hat{\mu}_1 + \frac{s_1 \sqrt{\varepsilon_1}}{\sqrt{T}}$ *would have different signs, and* $\text{sign}(\mu_1)$ *is not known.*

The condition (3.18) suggests that if $\varepsilon_1 < \left(\frac{\hat{\mu}_1}{s_1}\right)^2 T$, then $\text{sign}(\hat{\mu}_1) = \text{sign}(\mu_1)$. In other words, investors have no knowledge about the sign of μ_1 if the ambiguity parameter ε_1 rises above $\left(\frac{\hat{\mu}_1}{s_1}\right)^2 T$. The sign of μ_1 is especially important in the one-stock case or in the uncorrelated multi-stock cases because investors, intuitively, would not invest any wealth in stocks that have negative returns, and vice versa. Thus, if the ambiguity parameter ε_1 is such that the sign of μ_1 is unknown, investors would simply not invest in the risky asset.

Theorem 3.2.1 *If we solve the inner minimization problem in* (3.16) *with* (3.17), (3.16) *is equivalent to*

$$\max_{\omega_1} \left[\omega_1 \mu_1 - \frac{\gamma}{2} \omega_1^2 \left(\frac{s_1^2}{1 - \delta_1^l} \right) - |\omega_1| \frac{s_1 \sqrt{\varepsilon_1}}{\sqrt{T}} \right] \tag{3.19}$$

17

The optimization problem (3.19) *is maximized when*

$$
\omega_1 = \begin{cases}
\dfrac{\hat{\mu}_1 - \frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}}}{\gamma s_1^2}\left(1 - \delta_1^l\right) & , \quad \mu_1 > 0 \\[3ex]
0 & , \quad \varepsilon_1 > \left(\frac{\hat{\mu}_1}{s_1}\right)^2 T \\[3ex]
\dfrac{\hat{\mu}_1 + \frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}}}{\gamma s_1^2}\left(1 - \delta_1^l\right) & , \quad \mu_1 < 0
\end{cases}
\tag{3.20}
$$

Proof. We first assume that $\omega_1 \neq 0$. Otherwise, the portfolio is equivalent to a single risk-free asset, and in such case, there is no diversification involved. We define

$$
f\left(\omega_1, \mu_1\sigma_1\right) = \omega_1\mu_1 - \frac{\gamma}{2}\omega_1^2\sigma_1^2
$$

so the objective function (3.16) can be written simply as

$$
\max_{\omega_1}\left[\min_{\mu_1,\sigma_1} f\left(\omega_1, \mu_1, \sigma_1\right)\right]
$$

We first consider the inner minimization. The term $\frac{\gamma}{2}\omega_1^2\sigma_1^2$ is always non-negative regardless of ω_1, so the largest value possible for σ_1^2 would minimize $f\left(\omega_1, \mu_1, \sigma_1\right)$. The term $\omega_1\mu_1$ is a linear function of μ_1, so the value of μ_1 that minimizes $f\left(\omega_1, \mu_1, \sigma_1\right)$ depends on the sign of ω_1. If $\omega > 0$ ($\omega < 0$), then the smallest (largest) possible value of μ_1 would minimize $f\left(\omega_1, \mu_1, \sigma_1\right)$. Thus, if we solve the inner minimization problem, (3.16) becomes

$$
\max_{\omega_1}\left[\left(\hat{\mu}_1\omega_1 - \text{sign}\left(\omega_1\right)\frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}}\right) - \frac{\gamma}{2}\omega_1^2\left(\frac{s_1^2}{1 - \delta_1^l}\right)\right]
$$

With $\omega \cdot \text{sign}\left(\omega_1\right) = |\omega_1|$, it is equivalent to

$$
\max_{\omega_1}\left[\omega_1\hat{\mu}_1 - |\omega_1|\frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}}\hat{\mu}_1 - \frac{\gamma s_1^2\omega_1^2}{2(1 - \delta_1^l)}\right]
$$

Define $g\left(\omega_1\right) = \omega_1\hat{\mu}_1 - |\omega_1|\frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}} - \frac{\gamma s_1^2\omega_1^2}{2(1-\delta_1^l)}$. Since $|\omega_1|$ is differentiable at $\omega_1 \neq 0$, and we assumed $\omega_1 \neq 0$,

$$
\frac{\partial g\left(\omega_1\right)}{\partial \omega_1} = \hat{\mu}_1 - \frac{\omega_1}{|\omega_1|} \cdot \frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}} - \frac{\gamma s_1^2\omega_1}{1 - \delta_1^l}
$$

18

From Remark 3.2.1, we know that $\omega_1 = 0$ if $\varepsilon_1 > \left(\frac{\hat{\mu}_1}{s_1}\right)^2 T$, so we assume $\varepsilon_1 < \left(\frac{\hat{\mu}_1}{s_1}\right)^2 T$, which means $\text{sign}(\mu_1) = \text{sign}(\hat{\mu}_1)$. In one-stock case, we can easily verify that $\text{sign}(\omega_1) = \text{sign}(\hat{\mu}_1)$. Then, the first order condition above is equivalent to

$$\frac{\partial g(\omega_1)}{\partial \omega_1} = \begin{cases} \hat{\mu}_1 - \dfrac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}} - \dfrac{\gamma s_1^2 \omega_1}{1 - \delta_1^l} & , \quad \hat{\mu}_1 > 0 \\[3mm] \hat{\mu}_1 + \dfrac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}} - \dfrac{\gamma s_1^2 \omega_1}{1 - \delta_1^l} & , \quad \hat{\mu}_1 < 0 \end{cases} \tag{3.21}$$

By setting $\frac{\partial g(\omega_1)}{\partial \omega_1} = 0$, we obtain

$$\omega_1 = \begin{cases} \dfrac{\hat{\mu}_1 - \frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}}}{\gamma s_1^2} & , \quad \hat{\mu}_1 > 0 \\[5mm] \dfrac{\hat{\mu}_1 + \frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}}}{\gamma s_1^2} & , \quad \hat{\mu}_1 < 0 \end{cases} \tag{3.22}$$

as desired. Since
$$\frac{\partial^2 g(\omega_1)}{\partial \omega_1^2} = -\frac{\gamma s_1^2}{1 - \delta_1^l} < 0$$
the solution (3.20) indeed maximizes $g(\omega_1)$. $\qquad\square$

The solution (3.20) has several implications. First, when perfect information is available to investors, or equivalently, $\varepsilon_1 = 0$ and $\delta_1^l = 0$, the solution agrees with the solution to the classical mean-variance portfolio model given in (2.2). Second, as uncertainty increases, mathematically represented by a rise in ε_1 and δ_1^l, ω_1 decreases. The less confident investors are about the true parameters, the less they invest in the risky asset. Third, if the mean ambiguity parameter ε_1 rises above a certain level, namely $\left(\frac{\hat{\mu}_1}{s_1}\right)^2 T$, no wealth is invested in the risky asset. In plain words, it shows that too much uncertainty is too risky. Lastly, uncertainty in the variance of the risky asset does not change the investment plan (i.e. short or long) whereas uncertainty in the mean return of the risky asset certainly does.

Chapter 4

Generalization: Two-Stock Case

The one-stock case illustrates how we can model ambiguity in mean returns and variances. In reality, a portfolio usually contains more than one risky asset, and correlation coefficients among the risky assets should also be considered. On that note, we now generalize the idea presented in the earlier section to multi-assets, and we consider a portfolio with two risky assets to illustrate the generalization.

In the first section of the chapter, we first study the distribution of the sample correlation coefficient. The distribution is used to impose bounds on the correlation coefficient at a prescribed confidence level. We also discuss the distribution and the hypothesis testing of the sample covariance matrix. Then, we consider three scenarios. First, we analyze the case of unknown variances, assuming that the correlation coefficient can be perfectly estimated. The assumption here is somewhat unrealistic, but the objective of the first scenario is to investigate how ambiguity in variances may affect portfolio rules with a given correlation. Second, we study the opposite case of unknown correlation coefficient, assuming that the variances of the risky assets are perfectly estimated by sample variances. Lastly, we study the case of unknown covariance matrix as a whole, which is the most realistic scenario in practice.

4.1 Two-Stock Case: No Short-Selling

The main difference of the two stock-case from the one-stock case is the presence of the correlation coefficient between the risky assets. From the standard mean-variance model (2.1), the inner minimization illustrated in (3.11) for the one-stock case is now with respect to both mean returns, μ, and the entire covariance matrix, Σ. Then, the model is formulated as

$$\max_{\omega} \left[\min_{\mu, \Sigma} \left[\omega^T \mu - \frac{\gamma}{2} \omega^T \Sigma \omega \right] \right] \tag{4.1}$$

subject to

$$\mu \in M \subset \mathbb{R}^{2 \times 1} \ , \ \Sigma \in S \subset \mathbb{R}^{2 \times 2} \tag{4.2}$$

where M and S are determined based on a prescribed level of significance. In the two-stock case, if we hold one of the parameters certain, we can simplify (4.1), and analyze how portfolio rules change with respect to one parameter when the other is held certain (no ambiguity). In this section, we impose the restriction of no short sale, or $\omega \geq 0$. We first consider simpler cases where ambiguity is assumed on only one of variance and correlation with the other being held known.

4.1.1 Two-Stock Case: Unknown Variances

The objective is to solve the max-min problem (4.1). If we simplify (4.1) in linear form, the problem is equivalent to

$$\max_{\omega_1, \omega_2} \left[\min_{\mu_1, \mu_2, \sigma_1, \sigma_2, \rho} \left[(\omega_1 \mu_1 + \omega_2 \mu_2) - \frac{\gamma}{2} \left(\omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2 \omega_1 \omega_2 \beta \sigma_1 \sigma_2 \right) \right] \right] \tag{4.3}$$

If we impose the restriction of no short sale, or in other words, $\omega_1, \omega_2 \geq 0$, then solving the optimization problem becomes rather straightforward. We suppose that it is the case. We also note that the correlation coefficient between the two risky assets is the sample

correlation coefficient, β, as we assume that it is perfectly estimated by sample. The conditions (4.2) can be restated as

$$
\begin{array}{ccccc}
\hat{\mu}_i - \dfrac{s_i\sqrt{\varepsilon_i}}{\sqrt{T}} & \leq & \mu_i & \leq & \hat{\mu}_i + \dfrac{s_i\sqrt{\varepsilon_i}}{\sqrt{T}} \\[2ex]
\dfrac{s_i^2}{1 + \delta_i^u} & \leq & \sigma_i^2 & \leq & \dfrac{s_i^2}{1 - \delta_i^l}
\end{array}
\tag{4.4}
$$

Proposition 4.1.1 *If we solve 4.3 for ω_1 and ω_2, with 4.4, we obtain*

$$
\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \frac{1}{\gamma}
\begin{bmatrix} (\sigma_1^u)^2 & \beta\sigma_1^u\sigma_2^u \\ \beta\sigma_1^u\sigma_2^u & (\sigma_2^u)^2 \end{bmatrix}^{-1}
\begin{bmatrix} \mu_1^l \\ \mu_2^l \end{bmatrix}
\tag{4.5}
$$

or in explicit form,

$$
\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \frac{1}{\gamma}
\begin{bmatrix}
\dfrac{(\sigma_2^u)^2\,\mu_1^l - \beta\sigma_1^u\sigma_2^u\mu_2^l}{(\sigma_1^u)^2\,(\sigma_2^u)^2\,(1 - \rho_u^2)} \\[3ex]
\dfrac{(\sigma_1^u)^2\,\mu_2^l - \beta\sigma_1^u\sigma_2^u\mu_1^l}{(\sigma_1^u)^2\,(\sigma_2^u)^2\,(1 - \rho_u^2)}
\end{bmatrix}
\tag{4.6}
$$

where

$$
\begin{array}{ccccccc}
\mu_i^l & = & \hat{\mu}_i - \dfrac{s_i\sqrt{\varepsilon_i}}{\sqrt{T}} & , & \mu_i^u & = & \hat{\mu}_i + \dfrac{s_i\sqrt{\varepsilon_i}}{\sqrt{T}} \\[2ex]
\left(\sigma_i^l\right)^2 & = & \dfrac{s_i^2}{1 + \delta_i^u} & , & \left(\sigma_i^u\right)^2 & = & \dfrac{s_i^2}{1 - \delta_i^l}
\end{array}
\tag{4.7}
$$

Proof. The inner minimization problem is with respect to mean returns μ_1, μ_2 and covariance matrix σ_1, σ_2. We define

$$
f\left(\mu_1, \mu_2, \sigma_1, \sigma_2\right) = f_1\left(\mu_1, \mu_2\right) - f_2\left(\sigma_1, \sigma_2\right)
\tag{4.8}
$$

where

$$
\begin{array}{ccl}
f_1\left(\mu_1, \mu_2\right) & = & \omega_1\mu_1 + \omega_2\mu_2 \\[1ex]
f_2\left(\sigma_1, \sigma_2\right) & = & \dfrac{\gamma}{2}\left(\omega_1^2\sigma_1^2 + \omega_2^2\sigma_2^2 + 2\omega_1\omega_2\beta\sigma_1\sigma_2\right)
\end{array}
\begin{array}{r}
(4.9) \\[2ex]
(4.10)
\end{array}
$$

22

Then, the problem is equivalent to minimizing $f(\mu_1, \mu_2, \sigma_1, \sigma_2)$, which is then equivalent to minimizing $f_1(\mu_1, \mu_2)$ with respect to μ_1, μ_2 and maximize $f_2(\sigma_1, \sigma_2)$ with respect to σ_1, σ_2. Here, $f_1(\mu_1, \mu_2)$ is a linear function of μ_1 and μ_2, so it is minimized at the lowest values of μ_1 and μ_2. The max-min problem (4.3) becomes

$$\max_{\omega_1, \omega_2} \left[\left(\omega_1 \mu_1^l + \omega_2 \mu_2^l \right) - \frac{\gamma}{2} \left(\omega_1^2 \left(\sigma_1^u \right)^2 + \omega_2^2 \left(\sigma_2^u \right)^2 + 2 \omega_1 \omega_2 \beta \sigma_1^u \sigma_2^u \right) \right] \tag{4.11}$$

We define $g(\omega_1, \omega_2)$ to be the function in (4.11). The first order conditions are

$$\frac{\partial g(\omega_1, \omega_2)}{\partial \omega_1} = \mu_1^l - \gamma \left(\omega_1 \left(\sigma_1^u \right)^2 + \omega_2 \beta \sigma_1^u \sigma_2^u \right) \tag{4.12}$$

$$\frac{\partial g(\omega_1, \omega_2)}{\partial \omega_2} = \mu_2^l - \gamma \left(\omega_2 \left(\sigma_2^u \right)^2 + \omega_1 \beta \sigma_1^u \sigma_2^u \right) \tag{4.13}$$

We set the first order conditions to 0, and then solve for ω_1, ω_2 to obtain the solution. $\square$
The solution is valid only when

$$\left(\sigma_2^l \right)^2 \mu_1^l - \beta \sigma_1^u \sigma_2^u \mu_2^l \geq 0 \ , \ \left(\sigma_1^l \right)^2 \mu_2^l - \beta \sigma_1^u \sigma_2^u \mu_1^l \geq 0$$

Re-arranging the condition, we obtain

$$\beta \ \leq \ \frac{\sigma_2^u}{\sigma_1^u} \cdot \frac{\mu_1^l}{\mu_2^l} \tag{4.14}$$

This condition tells us that the solution is only valid when the correlation is smaller than the ratio of two Sharpe ratios. This result agrees with the fact that if the correlation is high, there is no (or less) merit in diversification, and hence, the proportion invested in one of the risky asset becomes zero (or minimal). Mathematically, if the correlation is higher than $\frac{\sigma_2^u}{\sigma_1^u} \cdot \frac{\mu_1^l}{\mu_2^l}$, no diversification is needed, and the entire wealth would be invested in the stock with a higher rate of return. We next consider the other case where we assume that the variances of the risky assets are perfectly estimated by sample, and allow ambiguity in the correlation between the assets.

4.1.2   Case of Unknown Correlation

In this section, we consider the case where the variances of the risky assets are assumed to be perfectly estimated by the sample variances, and mean returns and correlation coefficient

of the stocks leave room for uncertainty. Then, the problem (4.1) is simplifed, in a linear form,

$$
\max_{\omega_1,\omega_2} \left[\min_{\mu_1,\mu_2,\rho} \left[\omega_1\mu_1 + \omega_2\mu_2 - \frac{\gamma}{2} \left(\omega_1^2 s_1^2 + \omega_2^2 s_2^2 + 2\omega_1\omega_2\rho s_1 s_2 \right) \right] \right] \tag{4.15}
$$

subject to

$$
\hat{\mu}_i - \frac{s_i\sqrt{\varepsilon_i}}{\sqrt{T}} \leq \mu_i \leq \hat{\mu}_i + \frac{s_i\sqrt{\varepsilon_i}}{\sqrt{T}} \; , \; i = 1,2 \tag{4.16}
$$

$$
\rho_l \leq \rho \leq \rho_u
$$

Definition 4.1.1 *We define the adjusted Sharpe ratio, θ, as*

$$
\theta = \frac{\hat{\mu} - \frac{s\sqrt{\varepsilon}}{\sqrt{T}}}{s} \tag{4.17}
$$

The adjusted Sharpe ratio resembles the actual Sharpe ratio, but it takes the mean ambiguity into account, with the standard deviation estimated by its sample counterpart.

Proposition 4.1.2 *With $\omega \geq 0$, the solution to (4.15) subject to (4.16) is*

$$
\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \frac{1}{\gamma} \begin{bmatrix} s_1^2 & \rho_u s_1 s_2 \\ \rho_u s_1 s_2 & s_2^2 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mu}_1 - \frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}} \\ \hat{\mu}_2 - \frac{s_2\sqrt{\varepsilon_2}}{\sqrt{T}} \end{bmatrix} \tag{4.18}
$$

or equivalently

$$
\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \begin{bmatrix} \frac{\hat{\mu}_1 - \frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}}}{(1-\rho_u^2)s_1^2} - \frac{\hat{\mu}_2 - \frac{s_2\sqrt{\varepsilon_2}}{\sqrt{T}}}{(1-\rho_u^2)s_1 s_2}\rho_u \\[4mm] \frac{\hat{\mu}_2 - \frac{s_2\sqrt{\varepsilon_2}}{\sqrt{T}}}{(1-\rho_u^2)s_2^2} - \frac{\hat{\mu}_1 - \frac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}}}{(1-\rho_u^2)s_1 s_2}\rho_u \end{bmatrix} \tag{4.19}
$$

Proof. Following the proof of Theorem 3.2.1, we solve the inner minimization of (4.15) to obtain

$$
\max_{\omega_1,\omega_2} \left[\omega_1\mu_1^l + \omega_2\mu_2^l - \frac{\gamma}{2} \left(\omega_1^2 s_1^2 + \omega_2^2 s_2^2 + 2\omega_1\omega_2\rho_u s_1 s_2 \right) \right] \tag{4.20}
$$

24

where $\mu_i^l = \hat{\mu}_i - \frac{s_i \sqrt{\varepsilon_i}}{\sqrt{T}}$. We define

$$f(\omega_1, \omega_2) = \omega_1 \mu_1^l + \omega_2 \mu_2^l - \frac{\gamma}{2}\left(\omega_1^2 s_1^2 + \omega_2^2 s_2^2 + 2\omega_1 \omega_2 \rho_u s_1 s_2\right) \qquad (4.21)$$

The optimization (4.20) is two-dimensional, and we obtain the first order conditions as

$$\frac{\partial f(\omega_1, \omega_2)}{\partial \omega_1} = \mu_1^l - \gamma\left(\omega_1 s_1^2 + \omega_2 r_u s_1 s_2\right) \qquad (4.22)$$

$$\frac{\partial f(\omega_1, \omega_2)}{\partial \omega_2} = \mu_2^l - \gamma\left(\omega_2 s_2^2 + \omega_1 r_u s_1 s_2\right) \qquad (4.23)$$

Setting the first order conditions to 0 and rearranging,

$$\begin{aligned} s_1^2 \omega_1 + \rho_u s_1 s_2 \omega_2 &= \frac{\mu_1^l}{\gamma} \\ \rho_u s_1 s_2 \omega_1 + s_2^2 \omega_2 &= \frac{\mu_2^l}{\gamma} \end{aligned} \qquad (4.24)$$

If we write (4.24) into matrix notations, we obtain (4.18). For the solution (4.18) to be valid, we have to check that it is non-negative. Using Definition 4.1.1, the solution we obtain is

$$\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \frac{1}{\gamma} \begin{bmatrix} \dfrac{\theta_1 - \theta_2 \rho_u}{s_1(1 - \rho_u^2)} \\[3mm] \dfrac{\theta_2 - \theta_1 \rho_u}{s_2(1 - \rho_u^2)} \end{bmatrix} \qquad (4.25)$$

From (4.25), the solution is valid when

$$\theta_1 - \theta_2 \rho_u \geq 0 \ , \ \theta_2 - \theta_1 \rho_u \geq 0$$

or equivalently,

$$\min\left(\frac{\theta_1}{\theta_2}, \frac{\theta_2}{\theta_1}\right) \geq \rho_u \qquad (4.26)$$

$\square$

The condition (4.26) implies that the ratio of the adjusted Sharpe ratios between the two risky assets should be at least greater than ρ_u. Otherwise, no wealth is invested in the stock with the smaller adjusted Sharpe ratio. In other words, if one adjusted Sharpe ratio is relatively small compared to the other, there is no merit in investing in the stock with the smaller adjusted Sharpe ratio *if* no short sale is allowed.

25

Proposition 4.1.3 *If no short sale is allowed, and uncertainty in the correlation coefficient between the two stocks increases, then a proportion of wealth decreases in the stock with the smaller adjusted Sharpe ratio.*

Proof. Without loss of generality, suppose $\theta_1 \leq \theta_2$. Re-writing (4.19),

$$\omega = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \frac{1}{\gamma} \begin{bmatrix} \dfrac{\theta_1 - \theta_2 \rho_u}{s_1(1 - \rho_u^2)} \\[3mm] \dfrac{\theta_2 - \theta_1 \rho_u}{s_2(1 - \rho_u^2)} \end{bmatrix} \tag{4.27}$$

Then,

$$\frac{\partial \omega}{\partial \rho_u} = \frac{1}{\gamma} \begin{bmatrix} \dfrac{2\theta_1 r_u - \theta_2 r_u^2 - \theta_2}{s_1(1 - \rho_u^2)^2} \\[3mm] \dfrac{2\theta_2 r_u - \theta_1 r_u^2 - \theta_1}{s_2(1 - \rho_u^2)^2} \end{bmatrix} \tag{4.28}$$

Note that

$$
\begin{aligned}
2\theta_1 \rho_u - \theta_2 \rho_u^2 - \theta_2 &= = -\theta_2 \left(\rho_u^2 - \frac{2\theta_2}{\theta_1} \rho_u \right) - \theta_2 \\
&= -\theta_2 \left(\rho_u - \frac{\theta_1}{\theta_2} \right)^2 \underbrace{-\theta_2 + \frac{\theta_1^2}{\theta_2}}_{\leq 0} \\
&\leq 0
\end{aligned}
$$

Since $s_1(1 - \rho_u^2)^2 \geq 0$, the inequality above implies that

$$\frac{\partial \omega_1(\rho_*)}{\partial \rho_u} \leq 0 \tag{4.29}$$

so that ω_1 and ρ_u move in the opposite direction. $\qquad\square$

We can also derive the solution for the case of no risk-free asset (all wealth is distributed among the risky assets only) as the corollary below shows.

Corollary 4.1.1 *With no risk-free asset, the additional constraint $\omega^T \cdot \mathbb{1} = 1$ is imposed, and the solution to (4.15) is*

$$\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \begin{bmatrix} \dfrac{(\mu_1^l - \mu_2^l) - \gamma(\rho_u s_1 s_2 - s_2^2)}{\gamma(s_1^2 + s_2^2 - 2\rho_u s_1 s_2)} \\[3em] \dfrac{(\mu_2^l - \mu_1^l) - \gamma(\rho_u s_1 s_2 - s_1^2)}{\gamma(s_1^2 + s_2^2 - 2\rho_u s_1 s_2)} \end{bmatrix} \tag{4.30}$$

Proof. This result directly follows from (2.3). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Now we study the distribution of the correlation coefficient, and explain how we obtained ρ_u and ρ_l defined earlier in the section. We define X_{ik} to be k^{th} observation of the ith stock. In two stock cases with T observations, $i = 1, 2$ and $j = 1, 2, \cdots, T$. Then, the sample correlation coefficient between stock i and stock j, r_{ij}, is defined as

$$r_{ij} = \frac{\sum_{k=1}^{T} \left(X_{ik} - \bar{X}_i\right)\left(X_{jk} - \bar{X}_j\right)}{\sqrt{\sum_{k=1}^{T} \left(X_{ik} - \bar{X}_i\right)^2}\sqrt{\sum_{k=1}^{T} \left(X_{jk} - \bar{X}_j\right)^2}} \tag{4.31}$$

where $\bar{X}_i = \frac{1}{T}\sum_{k=1}^{T} X_{ik}$ is the sample mean of stock i. We denote ρ_{ij} to be the population correlation coefficient between stock i and stock j. In two-stock cases, we simplify the notation as $r_{ij} = r$, $\rho_{ij} = \rho$. The distribution of a sample correlation coefficient r, proposed in [Hotelling (1953)], is given as

$$f(r; T, \rho) = \frac{T-2}{\sqrt{2\pi}} \cdot \frac{\Gamma(T-1)}{\Gamma\left(T-\frac{1}{2}\right)} \cdot \frac{(1-\rho^2)^{\frac{T-2}{2}}(1-r^2)^{\frac{T-1}{4}}}{(1-\rho r)^{T+\frac{1}{2}}} \, {}_2F_1\left(\frac{1}{2}, \frac{1}{2}; T-\frac{1}{2}; \frac{1+\rho r}{2}\right) \tag{4.32}$$

where T is the sample size, ρ is the correlation coefficient, and ${}_2F_1(a, b; c; z)$ is a hypergeometric function

$$_2F_1(a, b; c; z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \cdot \frac{z^n}{n!}, \quad (x)_n = \frac{\Gamma(x+n)}{\Gamma(x)}$$

We refer to [Anderson (1958)] for the derivation of the distribution (4.32). The cumulative distribution function $F(r_0) = P(r \le r_0)$ can then be obtained by

$$F(r_0; T, \rho) = \int_{-1}^{r_0} f(r)dr \tag{4.33}$$

27

Tables of values for (4.3.3) are provided in the appendix, and these numbers are verified with results from [David (1938)].

The main task is to construct bounds of the population correlation coefficient ρ at a certain confidence level. This can be achieved by hypothesis testing on the null hypothesis $H_0 : \rho = \rho_0$ against the alternative hypothesis $H_1 : \rho \neq \rho_0$. Based on [David (1938)], the region of rejection at the $\alpha\%$ level of significance is $[-1, r_l) \cup (r_u, 1]$ such that

$$[1 - F(r_u; T, \rho_0)] + F(r_l; T, \rho_0) = \alpha \tag{4.34}$$

where the function $F(\cdot, \cdot, \cdot)$ is defined in (4.33). It is often suggested that r_l and r_u are chosen so that

$$[1 - F(r_u; T, \rho_0)] = \frac{\alpha}{2} = F(r_l; T, \rho_0)$$

In other words, r_l and r_u are chosen in such a way that each tail probability is equal.

Example 4.1.1 *If we want to test the null hypothesis $H_0 : \rho = 0.3$ against the alternative hypothesis $H_1 : \rho \neq 0.3$ at the 5% level of significance with the sample size of 30, we should find r_l and r_u such that*

$$F(r_l; 30, 0.3) = 0.025 \ , \ \ F(r_u; 30, 0.3) = 0.975$$

and we find that $r_l = -0.0627$ and $r_u = 0.5998$. In simple plain words, at the 5% confidence level, ρ is inside $[r_l, r_u]$. Trivially, if we set $\alpha = 100\%$, $r_l = r_u$, implying that the bound is indeed equivalent to the point estimate for ρ and hence, no ambiguity.

Once we obtain r_l and r_u from (4.34), we can apply the same concept to the correlation coefficient as we did to the return and variance. If the level of significance is low, we note that a positive (negative) sample correlation does not guarantee that the true correlation will be positive (negative).

4.1.3 Case of Unknown Covariance Matrix

In practice, the variance and correlation coefficient of the risky assets are never perfectly estimated by sample, so it is natural to study the portfolio rule with ambiguity in both

variances and correlation coefficient. In other words, the covariance matrix as a whole is ambiguous. In the earlier section, we presented the confidence interval of the correlation coefficient. In this section, we assume that variances and correlation coefficient are independent, and we consider their bounds in isolation from one another. This naive assumption of independence would underestimate the portfolio rule as the width of the bound for the correlation coefficient with a given level of confidence is wider than that of the bound obtained from the covariance matrix as a whole.

The formulation of the problem is nearly identical to the previous case of unknown variance, but we revisit the problem with its derivation here. The objective is to solve

$$\max_{\omega_1,\omega_2} \left[\min_{\mu_1,\mu_2,\sigma_1,\sigma_2,\rho} (\omega_1\mu_1 + \omega_2\mu_2) - \frac{\gamma}{2} \left(\omega_1^2\sigma_1^2 + \omega_2^2\sigma_2^2 + 2\omega_1\omega_2\rho\sigma_1\sigma_2\right) \right] \qquad (4.35)$$

subject to

$$
\begin{aligned}
\hat{\mu}_i - \frac{s_i\sqrt{\varepsilon_i}}{\sqrt{T}} &\leq \mu_i \leq \hat{\mu}_i + \frac{s_i\sqrt{\varepsilon_i}}{\sqrt{T}} \\
\frac{s_i^2}{1+\delta_i^u} &\leq \sigma_i^2 \leq \frac{s_i^2}{1-\delta_i^l} \\
\rho_l &\leq \rho \leq \rho_u
\end{aligned}
\qquad (4.36)
$$

We define

$$
\begin{aligned}
f_1(\mu_1,\mu_2) &= \omega_1\mu_1 + \omega_2\mu_2 & (4.37) \\
f_2(\sigma_1,\sigma_2,\rho) &= \omega_1\sigma_1^2 + \omega_2^2\sigma_2^2 + 2\omega_1\omega_2\rho\sigma_1\sigma_2 & (4.38)
\end{aligned}
$$

The inner minimization of the problem (4.35) is equivalent to minimizing f_1 with respect to μ_1 and μ_2, and maximizing f_2 with respect to σ_1, σ_2, and ρ. Since we assume no short sale, and f_1 is a linear function of μ_1 and μ_2, the smallest possible values for μ_1 and μ_2 minimize f_1. Similarly, the largest value possible for ρ maximizes f_2. However, depending on the sign of ρ_u, values of σ_1 and σ_2 that maximize f_2 are different. If $\rho_u > 0$, then the largest possible values for σ_1 and σ_2 maximize f_2, and vice versa. We denote

$$
\begin{aligned}
\mu_i^l &= \hat{\mu}_i - \frac{s_i\sqrt{\varepsilon_i}}{\sqrt{T}} \quad , \quad \mu_i^u = \hat{\mu}_i + \frac{s_i\sqrt{\varepsilon_i}}{\sqrt{T}} \\
\sigma_i^{u2} &= \frac{s_i^2}{1-\delta_i^l} \quad , \quad \sigma_i^{l2} = \frac{s_i^2}{1+\delta_i^u}
\end{aligned}
\qquad (4.39)
$$

Then, (4.35) is equivalent to

$$\max_{\omega_1,\omega_2} \left[\left(\omega_1\mu_1^l + \omega_2\mu_2^l\right) - \frac{\gamma}{2}\left(\omega_1^2\sigma_1^{u2} + \omega_2^2\sigma_2^{u2} + 2\omega_1\omega_2\rho_u\sigma_1^*\sigma_2^*\right) \right] \tag{4.40}$$

where

$$\sigma_i^* = \begin{cases} \sigma_i^u & , & \rho_u > 0 \\ \sigma_i^l & , & \rho_u < 0 \end{cases} \tag{4.41}$$

If we solve the system for ω_1 and ω_2, we obtain

$$\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \frac{1}{\gamma} \begin{bmatrix} (\sigma_1^u)^2 & \rho_u\sigma_1^*\sigma_2^* \\ \rho_u\sigma_1^*\sigma_2^* & (\sigma_2^u)^2 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mu}_1 - \dfrac{s_1\sqrt{\varepsilon_1}}{\sqrt{T}} \\ \hat{\mu}_1 - \dfrac{s_2\sqrt{\varepsilon_2}}{\sqrt{T}} \end{bmatrix} \tag{4.42}$$

We can see that the solution resembles the solution for the case of unknown variance. In this section, we imposed ambiguity bounds on the variances and correlation independently. This approach does not solve the original problem accurately as the naive assumption of independence between variances and the correlation overestimates ambiguity. Therefore, the assumption would lead to a solution that is more conservative. Figure 4.1 shows how dependence between the variances and correlation would affect the solution plane. The numbers used in this illustration are $\rho = 0.4$ and $\sigma = 0.3$. The vertical axis represents the correlation, and the horizontal axis is the variance. In fact, the axes themselves also represent the bounds on each of the parameters. In other words, assuming the variance and correlation are independent, we would have the entire region of the figure as the solution plane. However, with dependence between the two parameters, the solution plane would actually be an oval as shown in the figure. We also note that the scatter plot shows more density around the centre of the oval.

In Section 4.3, we will discuss how we can incorporate dependence into our solution using the Wishart distribution. This distribution generalizes the χ^2 distribution to multiple dimensions.
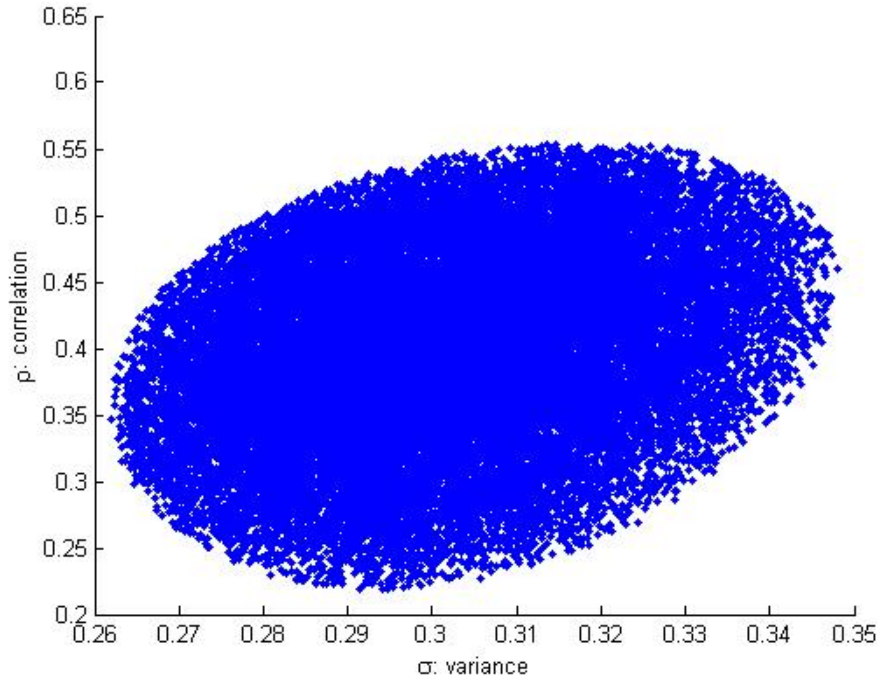
30

Figure 4.1: Dependence between Variance and Correlation

4.2 Two-Stock Case: Generalization

The ultimate objective is to obtain a general solution to the max-min problem (4.3). As presented in the earlier section, a certain restriction on ω_1, ω_2 leads to an analytic solution is available. However, for the general case, numerical methods may be required to obtain the solution, especially when the number of risky assets in the portfolio is large. The following two remarks are deemed useful to solve the max-min problem.

Remark 4.2.1 *We can transform the max-min problem into the minimax problem with the identity*

$$\max_x \min_{i \in I} f_i(x) = -\min_x \max_{i \in I} -f_i(x)$$

31

where I denotes the index set. In other words, $\{f_i\}_{i \in I}$ denotes a set of functions of x that we aim to obtain the maximum of the minimums of.

MATLAB provides a built-in numerical function to solve a minimax problem, so the identity in Remark 4.2.1 can be used to solve the max-min problem. The following remark illustrates how a double optimization problem could be simplified to a single optimization problem.

Remark 4.2.2 *For any $a, b \in \mathbb{R}$,*

$$\max(a, b) = \frac{a + b + |a - b|}{2} \tag{4.43}$$

This identity certainly holds for any real-valued functions. If, for $x \in \mathbb{R}$,

$$f_1(x) = 3x^2 - 2x + 1$$
$$f_2(x) = -x^2 + x - 4$$

then,

$$\min_x \left[\max_{i \in \{1,2\}} (f_1(x), f_2(x)) \right] = \min_x \left[\frac{(2x^2 - x - 3) + |4x^2 - 3x + 5|}{2} \right] \tag{4.44}$$

in which case the problem is simplified to a simple optimization problem in one dimension. A simple calculation gives $x = \dfrac{1}{3}$ as the solution to the above example.

In our context, x in Remark 4.2.2 is a vector of $\mu_1, \mu_2, \sigma_1, \sigma_2$, and ρ, and values of these parameters serve as coefficients of functions of ω_1 and ω_2. The double optimization problem faces the following complications:

1. The problem of our interest is not one dimensional.

2. The index set is not finite as $\mu_1, \mu_2, \sigma_1, \sigma_2$, and ρ can take any real values in intervals (4.4).

32

The first point can be handled by numerical methods for multi-dimensional optimization. One well-known method is Newton's method that involves calculations of the gradient, second derivatives, and the Hessian matrix with respect to the parameters. If we denote x_n to be the solution at n^{th} iteration, then the solution at the $n + 1^{th}$ iteration can be obtained by

$$x_{n+1} = x_n - \gamma \left[Hf\left(x_n\right) \right]^{-1} \triangledown f\left(x_n\right) \tag{4.45}$$

where H denotes the Hessian matrix, and $\triangledown$ is the gradient of the objective function $f\left(x\right)$, and $\gamma > 0$. We use approximations to handle the second point. We evenly divide intervals for each parameter in (4.4) to obtain finite number of possible values for each parameter. Then, we use numerical algorithm on the finite number of functions to solve the problem.

4.3   Case of Unknown Covariance Matrix

In the previous sections, we considered the notion of ambiguity for mean returns, variances, and the correlation separately by introducing ambiguity parameters for each parameter. Using the fact that the sample covariance matrix follows a Wishart distribution, we now study ambiguity aversion in the variance and correlation collectively. When returns of the stocks follow the multivariate normal distribution with the covariance matrix Σ, the sample covariance S follows a Wishart distribution $W\left(S; \Sigma, T - 1\right)$ where the probability density function is

$$p\left(S; \Sigma, T - 1\right) = \frac{s^{\frac{T-3}{2}}}{2^T |\Sigma|^{\frac{T}{2}} \Gamma_p\left(\frac{T}{2}\right)} \exp\left(-\frac{1}{2} \operatorname{tr}\left(\Sigma^{-1} S\right)\right) \tag{4.46}$$

We denote Σ_0 to be the true, unknown covariance, and S to be the sample covariance, and test the null hypothesis $H_0 : \Sigma = \Sigma_0$ against the alternative $H_1 : \Sigma \neq \Sigma_0$. Following notations and results presented in [Morrison (1967)], the test statistic L is

$$L = \nu \left(\ln |\Sigma_0| - \ln |S| + \operatorname{tr}\left(S\Sigma_0^{-1}\right) - p\right) \tag{4.47}$$

where $\nu = T - 1$ is a general degrees of freedom parameter, and p is the dimension ($p = 2$ for the two-stock case). For large T, L is distributed as a χ^2 variate with $\frac{1}{2}p\left(p + 1\right)$ degrees

of freedom. Based on L, [Barlett (1954)] introduced the scaled statistic

$$L' = \left[1 - \frac{1}{6(T-1)}\left(2p + 1 - \frac{2}{p+1}\right)\right]L \qquad (4.48)$$

and the decision rule is to reject H_0 if $L' > \chi^2_{\alpha;\frac{1}{2}p(p+1)}$. Since the objective is to minimize $\omega^T\mu - \frac{\gamma}{2}\omega^T\Sigma\omega$, we need to find Σ' that satisfies $L'(\Sigma') < \chi^2_{\alpha;\frac{1}{2}p(p+1)}$. We present the following numerical example to show how the Wishart distribution can be used in our problem.

Example 4.3.1 *Suppose we obtain the sample covariance matrix*

$$S = \begin{bmatrix} 0.6 & 0.2 \\ 0.2 & 0.4 \end{bmatrix}$$

from the sample size of $T = 27$. Our objective here is to find a matrix Σ' that follows $W(S,T)$ such that $L(\Sigma') < \chi^2_{\alpha;\frac{1}{2}p(p+1)}$ where we denote its elements as

$$\Sigma' = \begin{bmatrix} a_{11} & b \\ b & a_{22} \end{bmatrix}$$

In this example, suppose we use $\alpha = 0.5$. Since we are considering the two-stock case here, $p = 2$. From (4.48), we obtain

$$\begin{aligned} L(\Sigma') &= \left[1 - \frac{1}{6(T-1)}\left(2p + 1 - \frac{2}{p+1}\right)\right]L \\ &= \frac{35}{36}\left[26\left(\ln\left(a_{11}a_{22} - b^2\right) - \ln(0.2) + \frac{0.4a_{11} + 0.6a_{22} - 0.4b}{a_{11}a_{22} - b^2} - 2\right)\right] \end{aligned}$$

and this quantity should be less than $\chi^2_{0.5,3} = 25.3365$. Re-arranging the inequality,

$$\ln\left(a_{11}a_{22} - b^2\right) + \frac{0.4a_{11} + 0.6a_{22} - 0.4b}{a_{11}a_{22} - b^2} < 1.3929$$

In the inequality above, we see the dependence between the variances, a_{11} and a_{22}, and correlation, $\dfrac{b}{\sqrt{a_{11}a_{22}}}$. Our ultimate goal here is to find a matrix Σ' that maximizes $\frac{\gamma}{2}\omega^T\Sigma'\omega$ whose elements satisfy the inequality above.

34

One problem is that a_{11}, a_{22}, and/or b can still be far off from the sample covariance matrix even with a high level of confidence. Even with an additional condition $a_{11}a_{22} - b^2 > 0$ since a covariance matrix is positive-definite, we have 3 unknowns with 2 conditions which leaves the solution space for a_{11}, a_{22}, and b wide open. One solution to this problem is to use the bounds on the variances that have been discussed in the previous sections. It is widely known that variances are easier to estimate than mean returns and correlation coefficients. Additionally, variances are known to follow a well-known distribution so that they are easy to manipulate. Therefore, we impose the bounds on a_{11} and a_{22}, which in turn controls the bounds for b. That way, we ensure that with a certain level of confidence, the matrix we are looking for is not far off from the covariance matrix.

Chapter 5

Numerical Results

In this chapter, we apply results from previous chapters to numerical examples. Firstly, we use simulated returns, which are assumed to be normally distributed. This is easily done with a random number generator readily available in most programming languages, and we apply our results to these simulators. Next, we use actual returns calculated from indices as opposed to stocks. This is done to avoid any data issues, e.g. stock splits. Moreover, indices are more standard.

We us the Sharpe ratio to measure the performance of each portfolio. We use a different set of values for ambiguity parameters, and also vary the length of time frame to study the impact of ambiguity aversion. In the first section, we consider the simplest one-stock case. We compare results from the mean-variance portfolio, the portfolio with mean ambiguity (Garlappi, Uppal, Wang), and the portfolio with mean and variance ambiguity incorporated. In the second section, we consider the multi-asset case. We compare results from each portfolio to see the impact of each ambiguity parameter.

5.1 One-Stock Case

We define α to be the confidence level of the investors on the parameters. In our models, $\varepsilon = 0$ corresponds to no ambiguity in mean, and $\varepsilon = \infty$ corresponds to complete ambiguity in mean. The latter case means that there is absolutely no information available about the market. Ambiguity parameters in variance, δ^l, δ^u are defined in a similar fashion. We understand $\alpha = 0$ means that investors have no confidence in accuracy of the estimated parameters, and $\alpha = 1$ to mean perfect confidence in their estimation. The following tables summarizes values of $\varepsilon, \delta^l, \delta^u$ for given α.

α	ε	δ^l	δ^u
0.00	∞	1.0000	∞
0.05	1.9801	0.2378	0.2696
0.15	1.4489	0.1802	0.1922
0.25	1.1560	0.1469	0.1505
0.35	0.9383	0.1215	0.1201
0.45	0.7579	0.1000	0.0952
0.55	0.5995	0.0808	0.0737
0.65	0.4549	0.0631	0.0542
0.75	0.3194	0.0462	0.0361
0.85	0.1895	0.0298	0.0190
0.95	0.0628	0.0137	0.0025
1.00	0.0000	0.0000	0.0000

Table 5.1: Values of ambiguity parameters for given confidence level α, with $T = 60$

We see that when $\alpha = 0$, then $\varepsilon = \infty$, implying that the true mean return can possibly lie anywhere within $(-\infty, \infty)$. Likewise, $\alpha = 0$ leads to $\delta^l = 1$ and $\delta^u = \infty$. This implies that the ratio of the sample variance and true variance, $\dfrac{s^2}{\sigma^2}$, can possibly be any positive real number. For $\alpha = 1$, then $\varepsilon = 0, \delta_l = 0$, and $\delta^u = 0$, implying that confidence intervals degenerate into point estimates. In the next sections, we present the results of the one-stock case with simulated returns as well as empirical data.

5.1.1 Simulated Results

The numerical method of Monte Carlo simulations is a powerful tool, yet easy to implement, to verify the validity of the results if appropriate conditions are satisfied. Before we use empirical data, we use simulations as a quick check that the theory behind our model is indeed valid. Throughout this chapter, we use a window of $T = 60$ days to estimate the model parameters, and update the portfolio for the next 50 days. Based on the portfolio returns for the 50 days, we calculate the average return and the standard deviation, and eventually, the Sharpe ratio of the portfolio. We have prepared empirical data for S&P 500 and NIKKEI 225 indices. We extract 307 days of returns from these indices, from May 31, 2011 and back, calculate the mean return and standard deviation of each index, and treat those values as *true* parameter values for each index in our simulation. As we are using 60 observations to estimate the parameter to measure performance for the next 50 days, we simulate 110 returns. Table 5.2 outlines the values for mean returns and standard deviations we use in our simulations.

	S&P 500	NIKKEI 225
Mean	0.00495	-0.00233
St.Dev	0.00454	0.06238

Table 5.2: Mean daily returns and standard deviations for S&P 500 and NIKKEI 225

The objective here is to verify if ambiguity aversion leads to superior performance as we claimed it would in the earlier chapters. We consider simulating 110 observations as one run, and we test 100,000 runs to see how often the classical mean-variance model would outperform our model that incorporates ambiguity aversion. Based on our simulations, there is not one single run that the mean-variance portfolio outperformed the portfolio with ambiguity aversion, and no one single run that the portfolio with ambiguity in mean returns only outperformed the portfolio with ambiguity in both mean returns and the variance, with performance measured in terms of Sharpe ratios. Assuming that 100,000 runs are enough to make our case, we see that the notion of ambiguity certainly improves portfolio returns per unit risk. On rare instances depending on a set of normal random

numbers generated, Sharpe ratios for a few particular runs blow up due to extremely small standard deviations. We exclude those rare instances in from our calculations.

	S&P 500			NIKKEI 225		
	MV	A1	A2	MV	A1	A2
Average	0.1983	0.2352	0.2752	0.2741	0.3475	0.3670
St.Dev	0.00312	0.00409	0.00421	0.00193	0.00476	0.00492

Table 5.3: Average Sharpe ratios and standard deviations for 100,000 simulated runs

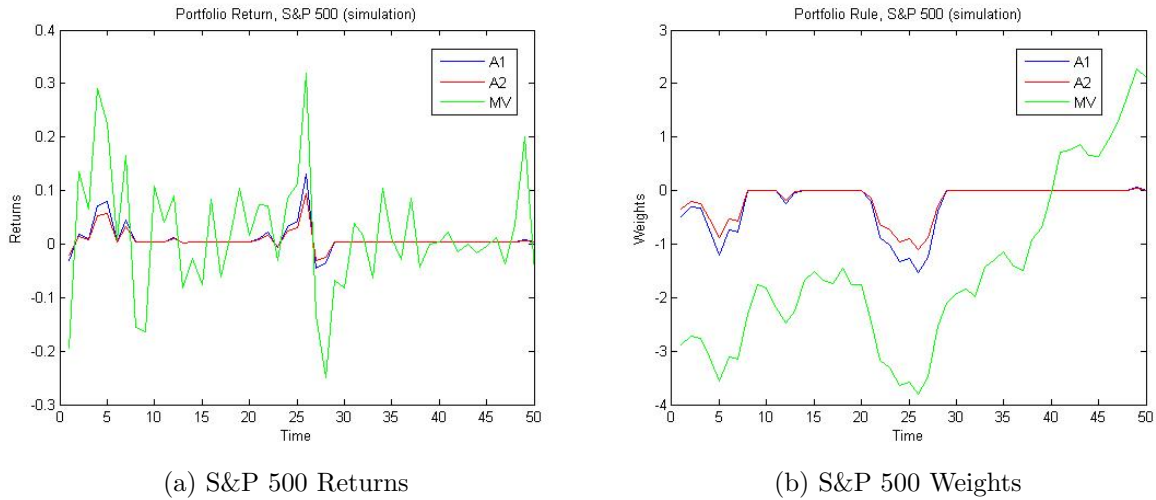(a) S&P 500 Returns   (b) S&P 500 Weights

Figure 5.1: One-Stock case, S&P 500 Simulation

MV refers to the classical mean-variance model, A1 is the portfolio with ambiguity in mean returns, and A2 is the portfolio with ambiguity in both mean returns and the variance. Numbers for A1 and A2 are based on $\alpha_\varepsilon = 0.50$ and $\alpha_\delta = 0.50$. We see that the performance of the portfolio improves as we incorporate more ambiguity into investment decisions, and as noted earlier, none of our runs has contradicted this result. Figure 5.1 and 5.2 show portfolio returns and weights over the 50 days for a single run 100,000 runs, as illustration purposes. We see that returns on the mean-variance portfolio is very volatile whereas ambiguity models reduces shocks substantially.
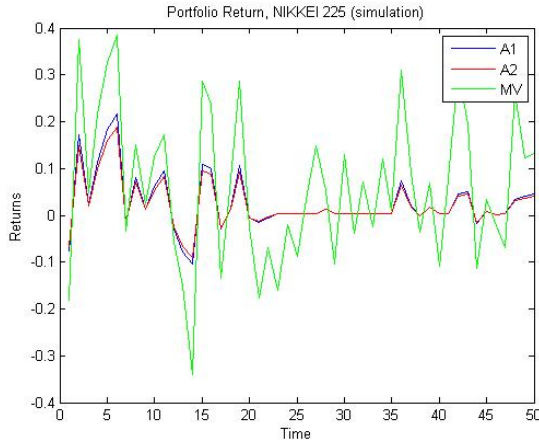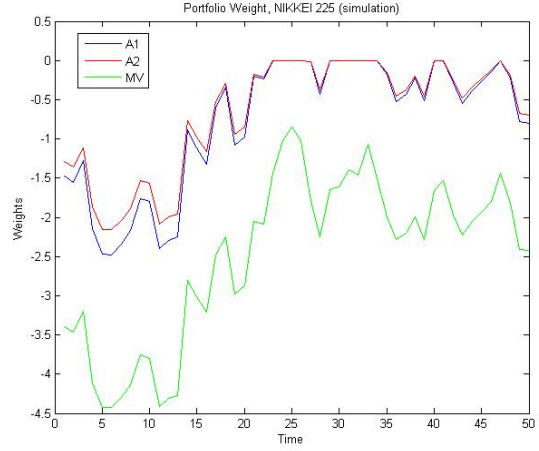
(a) NIKKEI 225 Returns   (b) NIKKEI 225 Weights

Figure 5.2: One-Stock case, NIKKEI 225 Simulation

5.1.2   Empirical Results

For the numerical results for one-stock case, we use the S&P 500 and NIKKEI 225 indices as our data. We use historical daily returns calculated by the difference between the two dates divided by the index value of the earlier date of the two. The daily risk-free rate is estimated to be 0.34% obtained from the 3-month bond yield by Bank of Canada. We denote T to be the estimation window, which means that if, for example, $T = 60$, we use 60 days of returns to estimate the parameters, to obtain $\hat{\mu}$ and s^2. We use T number of observations to study the performance of the portfolio for the next 50 days. In other words, if we use $T = 60$, then we use observations from Day 1 to Day 60 to estimate the parameters, which are used to determine the portfolio rule for Day 61. On Day 61, we use observations from Day 2 to Day 61 to estimate the parameters, which are then used to determine the portfolio rule for Day 61. We constantly update the portfolio rule with the most recent T number of observations for the next 50 days.

Table 5.4 and 5.5 shows Sharpe ratios for different combinations of ambiguity levels in mean returns and variance. Here, α_ε and α_δ denote the confidence level on the estimation

	α_δ					
	0.1	0.3	0.5	0.7	0.9	1.0
0.1	0.4053	0.3841	0.3741	0.3673	0.3619	0.3575
0.3	0.3752	0.3591	0.3515	0.3463	0.3422	0.3389
α_ε 0.5	0.3254	0.3123	0.3061	0.3019	0.2986	0.2959
0.7	0.2804	0.2696	0.2645	0.2610	0.2583	0.2561
0.9	0.2607	0.2521	0.2480	0.2480	0.2431	0.2413

Table 5.4: T = 60, Sharpe Ratio, S&P 500

	α_δ					
	0.1	0.3	0.5	0.7	0.9	1.0
0.3	1.7902	1.6088	1.5138	1.4440	1.3855	1.3586
0.5	0.5712	0.5228	0.4975	0.4789	0.4633	0.4561
α_ε 0.7	0.3846	0.3592	0.3459	0.3362	0.3280	0.3242
0.9	0.2335	0.2193	0.2118	0.2064	0.2018	0.1996

Table 5.5: T = 60, Sharpe Ratio, NIKKEI 225

of mean returns and variances, similar to α in Table 5.1. For example, $\alpha_{epsilon} = 0.5$ and $\alpha_\delta = 0.7$ means we set the confidence level on the estimation of mean returns at 50% and that of variances at 70%. Again, a higher α value translates into less ambiguity incorporated into the model, and vice versa. The mean-variance model with the same set of data gives the Sharpe ratio of 0.2375 for S&P 500 and 0.1229 for NIKKEI 225. The last column indicates the portfolio with ambiguity in mean returns only. These simulations are consistent with the main ideas of our result, namely:

1. The mean-variance model shows a high volatility in portfolio returns.

2. The weights on the indices are inflated for the mean-variance model, and ambiguity models are more conservative.

One important observation is that in ambiguity models, the weight on the risky asset
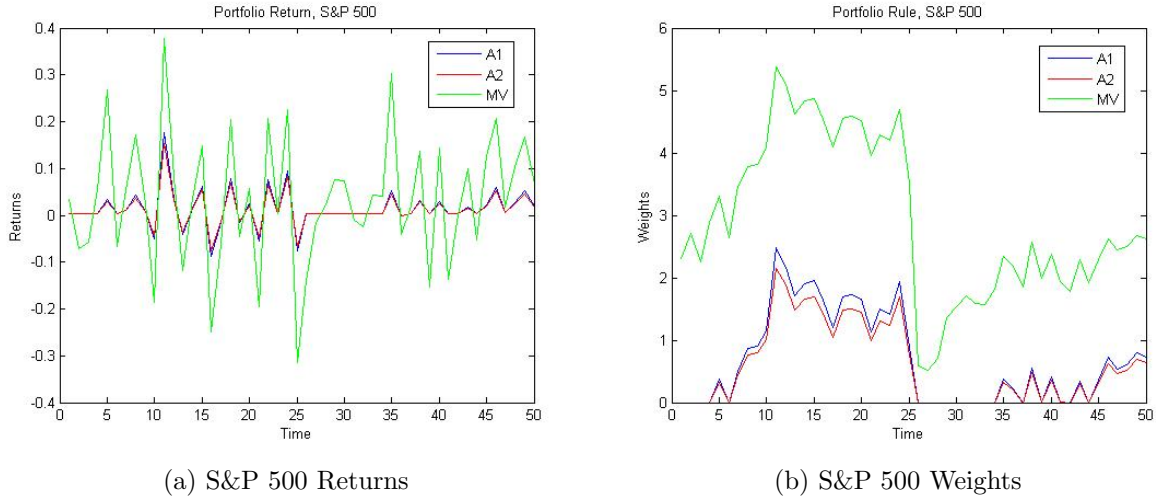
(a) S&P 500 Returns

(b) S&P 500 Weights

Figure 5.3: One-Stock case, S&P 500

(index in this case) could possibly be 0 for an extended period time. For the portfolio with S&P 500 index, for example, we see that the weight on the index is 0 from Day 26 to Day 33. This tells us that if the confidence level on parameter ambiguity is above a certain level, it is advisable not to invest in the risky asset.

5.2   Multi-Stock Case

We use a portfolio consisting of the S&P 500 and NIKKEI 225 indices, and a risk-free asset. The two-stock case is constructed in a similar way as the one-stock case. We picked the Japanese index since it has a somewhat smaller correlation with S&P 500 than any other major indices in North America. A smaller correlation would make a better case for diversification. We use $T = 60$ observations to construct a portfolio (or estimate the parameters) from March 16, 2011 to May 31, 2011, a total of 50 business days. We present the numerical results for each case for the two-asset model, outlined in Chapter 4.
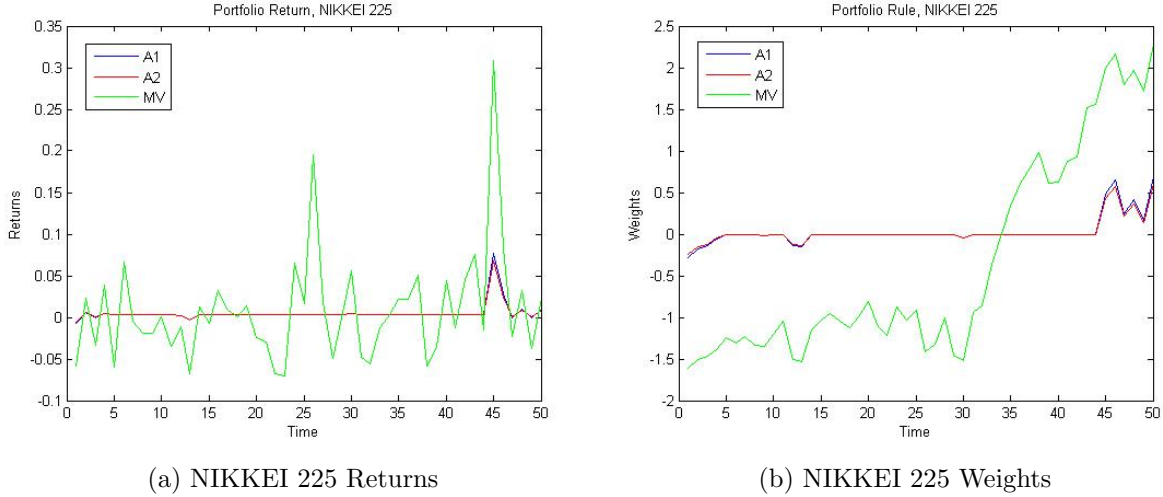
(a) NIKKEI 225 Returns   (b) NIKKEI 225 Weights

Figure 5.4: One-Stock case, NIKKEI 225

5.2.1 Two-Stock Model: Mean-Variance Model

We briefly present the results of the mean-variance model for comparison to our ambiguity models. The Sharpe ratio for the mean-variance model is 0.1342. Figure 5.5 shows the portfolio returns and weights on the indices based on the mean-variance model. We use these results to study how ambiguity impacts portfolio returns and rules.

5.2.2 Two-Stock Model: Unknown Variance

We first consider the model with ambiguity in mean returns and the variances. Assuming no ambiguity in correlation, we use the sample correlation between the two indices in the covariance matrix. As illustrated in Chapter 4, the variance of the risky asset follows the χ^2 distribution. Once we obtain the bounds on the variances, then we can use numerical approximations to find the optimal portfolio rules. The numerical algorithm is as follows: we divide the bounds of mean returns and variances into the finite number of sub intervals. Suppose we have $\sigma^2 \in [0.3, 0.8]$, for instance, and we divide this interval into $n = 5$ evenly
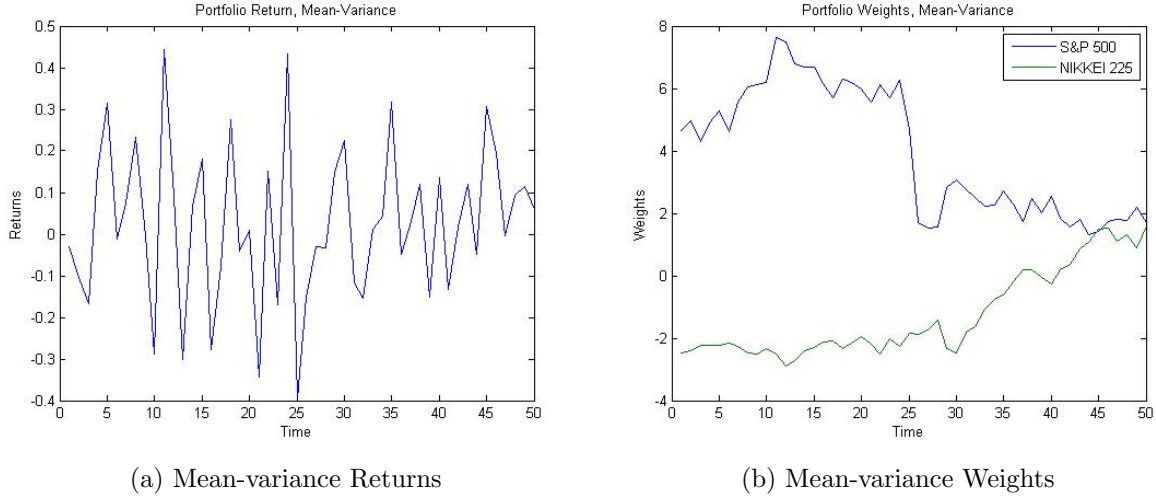
43

(a) Mean-variance Returns (b) Mean-variance Weights

Figure 5.5: Two-Stock Case, Mean-Variance Model

spaced sub intervals. Whether they are evenly divided or not becomes less important once n gets bigger. If we use $n = 5$, then we have 6 different values for σ^2, namely 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8. We do the same for mean returns. Then, there are a total of 36 combinations of values for the mean return and the variance. We use these combinations to obtain 36 different real-valued functions of ω. Then, we use the minimax solver to solve for ω, as presented in Section 4.2.

We use α_ε and α_δ between 0.3 and 0.8, and obtain Sharpe ratios for the portfolio with ambiguity in mean returns and variances based on these confidence levels. Table 5.6 shows Sharpe ratios for different levels of confidence in parameter estimates. Compared to the Sharpe ratio of the mean-variance model which is 0.1342, we notice that the performance of the portfolio has been greatly improved. We also note that to acknowledge less certainty in parameter estimates leads to higher Sharpe ratios. Compared to the Sharpe ratio at $\alpha_\varepsilon, \alpha_\delta = 0.3$ to the Sharpe ratio at $\alpha_\varepsilon, \alpha_\delta = 0.8$, it shows nearly 50% improvement. Lastly, one important result here is that a change in α_ε has a greater impact on the performance of the portfolio than α_δ. In other words, we see that ambiguity in mean returns plays a bigger role than in variances and this result is in line with the fact that mean returns are

| | α_δ | | | | | |
	0.3	0.4	0.5	0.6	0.7	0.8
0.3	0.4409	0.4324	0.4288	0.4234	0.4125	0.4111
0.4	0.4223	0.4166	0.4102	0.4044	0.3979	0.3933
0.5	0.4155	0.4062	0.3942	0.3830	0.3797	0.3782
α_ε 0.6	0.4029	0.3942	0.3812	0.3750	0.3691	0.3611
0.7	0.3854	0.3734	0.3648	0.3601	0.3529	0.3433
0.8	0.3701	0.3621	0.3544	0.3482	0.3410	0.3332

Table 5.6: T = 60, Sharpe ratios for two-stock portfolio with S&P 500 and NIKKEI 225, ambiguity in mean returns and variances

generally more difficult to correctly estimate than the variances.

Figure 5.6 shows portfolio returns and weights on S&P 500 and NIKKEI 225 indices, at $\alpha_\varepsilon, \alpha_\delta = 0.50$. We see that the ambiguity model advises substantially decreased weights on both indices, and as a result, the portfolio returns show less volatility or shocks, which in turn leads to higher Sharpe ratios. It is intuitive that weights on NIKKEI 225 are substantially smaller because the sample mean return on NIKKEI 225 is negative and smaller than that of S&P 500.

5.2.3   Two-Stock Model: Unknown Correlation

In this section, we consider another case where we assume ambiguity in the mean returns and correlation coefficient. The algorithm is very similar to the earlier case in Chapter 5.2.2. Compared to the mean returns and variances, correlation coefficients have not been the focus of research in portfolio theory. In this section, we consider the model with ambiguity in mean returns and the correlation coefficient, with variances estimated by the sample variances.

We notice that the impact of ambiguity in the correlation coefficient is not as large as the impact of ambiguity in variances. Table 5.7 shows values for Sharpe ratios for different
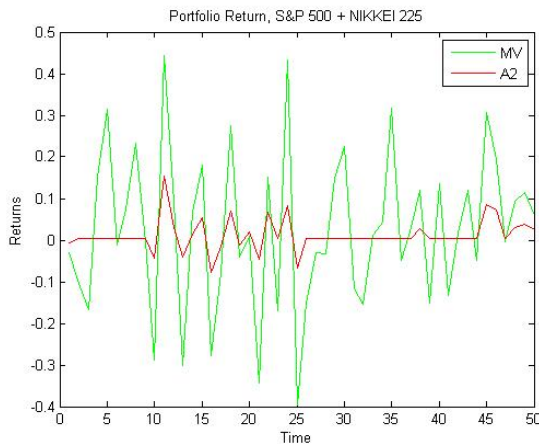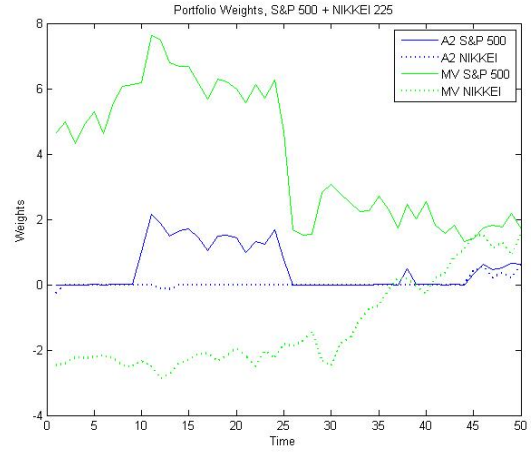
(a) A2 Returns   (b) A2 Weights

Figure 5.6: Two-Stock Case, Ambiguity in mean returns and variances

values of α_ε and α_ρ. Although ambiguity in correlation certainly improves the performance of the portfolio compared to the mean-variance model, the results indicate that the impact might be as large.

5.2.4 Two-Stock Model: Unknown Covariance Matrix

We now consider the case of the unknown covariance matrix. The numerical algorithm for ambiguity on the covariance matrix is explained in Chapter 4.3. In practice, estimations error is present in both the variances and correlation coefficient, and the result for this section would be the closest to practical situations.

The Sharpe ratio for the classical mean variance portfolio is 0.1342. In Table 5.8, α_ε denotes a confidence level on mean returns, and α_Σ on the covariance matrix. Table 5.8 summarizes Sharpe ratios for the portfolio with ambiguity in mean returns and covariance matrices. We see that Sharpe ratios for the portfolio with ambiguity are much higher than that of the classical mean-variance portfolio. It is also important to note that a change

	α_ρ					
	0.3	0.4	0.5	0.6	0.7	0.8
0.3	0.4415	0.4379	0.4301	0.4233	0.4187	0.4161
0.4	0.4297	0.4221	0.4159	0.4099	0.4023	0.3989
0.5	0.4159	0.4091	0.4026	0.3973	0.3900	0.3828
α_ε 0.6	0.4066	0.3929	0.3849	0.3777	0.3703	0.3649
0.7	0.3892	0.3810	0.3739	0.3661	0.3559	0.3474
0.8	0.3728	0.3644	0.3591	0.3504	0.3478	0.3421

Table 5.7: T = 60, Sharpe ratios for two-stock portfolio with S&P 500 and NIKKEI 225, ambiguity in mean returns and correlation coefficient

in confidence level on mean returns has a bigger impact on Sharpe ratios than a change in confidence level on the covariance matrix. As noted in one-stock case, a higher level of confidence tends to lead to smaller Sharpe ratios. This is mainly due to a higher variance of portfolio returns when we are less conservative.

5.3   Comparison

In the earlier sections, we presented the numerical results for different cases of ambiguity portfolios consisting of S&P 500 and NIKKEI 225 indices. The sample correlation coefficient between the two indices based on 307 observations is 0.4694. Since it is even below 0.50, the effect of diversification could be useful. As we compare Sharpe ratios for each case, we notice that introducing ambiguity in variances, correlation, or the covariance matrix as a whole improves the performance of the portfolio in terms of Sharpe ratio, as opposed to ambiguity in mean returns only. However, the impact from ambiguity in correlation coefficient is not as significant as that from ambiguity in variances or the covariance matrix as a whole. In addition, it is computationally more expensive to obtain bounds for correlation coefficients, than to obtain bounds for the variances or covariance matrix, as the distribution for the correlation coefficient is much more complex and not well-known.
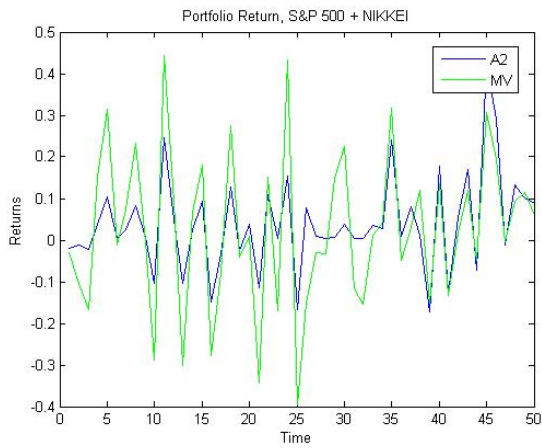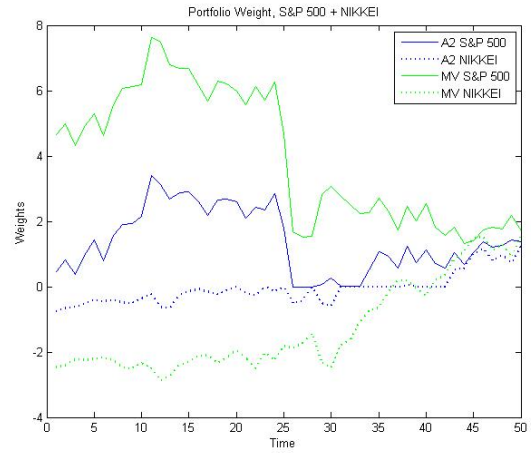
(a) A2 Returns   (b) A2 Weights

Figure 5.7: Two-Stock Case, Ambiguity in mean returns and correlation coefficient

For each case, it is advisable to short sell NIKKEI and buy long S&P 500. This is intuitive because the mean return on the S&P 500 index is positive, and negative on the NIKKEI 225 index.

		α_Σ					
		0.3	0.4	0.5	0.6	0.7	0.8
	0.3	0.5245	0.5172	0.5109	0.5042	0.4958	0.4892
	0.4	0.4989	0.4833	0.0.4787	0.4703	0.4656	0.4609
	0.5	0.4841	0.4774	0.4699	0.4620	0.4578	0.4519
α_ε	0.6	0.4687	0.4615	0.4579	0.4523	0.4469	0.4401
	0.7	0.4555	0.4498	0.4431	0.4381	0.4211	0.4098
	0.8	0.4391	0.4269	0.4153	0.4021	0.3879	0.3725

Table 5.8: T = 60, Sharpe ratios for two-stock portfolio with S&P 500 and NIKKEI 225, ambiguity in mean returns and covariance matrix

α_ε	0.3	0.4	0.5	0.6	0.7	0.8
Sharpe Ratio	0.3917	0.3822	0.3703	0.3525	0.3321	0.3264

Table 5.9: Sharpe ratios, Portfolio with ambiguity in mean returns

Chapter 6

Conclusion

Ever since mean-variance analysis has been proposed by Harry Markowitz in 1952, the importance of diversification has constantly been addressed. The theory is based on balancing the returns of the portfolio against the standard deviation of return, and is a great improvement over the naive expected value criterion. The main concept of the classical mean-variance portfolio model is that an investment is a tradeoff between risk and expected return where an asset's return is typically assumed to be normally distributed. Although the model has now been widely used in the industry as well as in academia, many researchers have challenged its limitations. The key limitations are the absence of estimation risk and the lack of the behavioral aspects of decision making. Using the notion of ambiguity aversion discussed in the literature, we use statistical distributions to incorporate this in the covariance matrix of the risky assets as well as their returns.

The basic idea of ambiguity aversion resembles that of stress testing as we are considering the minimum performance of the portfolio given specific ranges for the model parameters. In other words, we argue that rational investors are averse to ambiguity in parameters, and try to invest conservatively whenever such ambiguity is present. We measure the performance of the portfolio using its Sharpe ratio, and our results show that the model with ambiguity incorporated has higher Sharpe ratios.

The model and our analysis still leave room for improvement. We assumed that the daily stock returns are normally distributed, and we may support this assumption using the Central Limit Theorem, as briefly noted in Chapter 3. However, in practice, it is not uncommon to observe a standard deviation movement larger than, say, 3, which is exceptionally rare with the normal distribution. Therefore, it could also be useful to consider other appropriate distributions. In addition, although we are able to obtain closed-form solutions when we impose the no short sale restriction, we rely on simulations and numerical methods to solve the general case. With some reasonable restrictions, we might be able to derive closed-form solutions for the general case.

Appendix A

Cumulative Distribution Table

$\rho\backslash$ r	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0
-0.9	0.534	0.979	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.8	0.031	0.530	0.902	0.986	0.998	1.000	1.000	1.000	1.000	1.000
-0.7	0.001	0.126	0.526	0.837	0.958	0.991	0.998	1.000	1.000	1.000
-0.6	0.000	0.020	0.196	0.523	0.792	0.929	0.980	0.995	0.999	1.000
-0.5	0.000	0.003	0.054	0.240	0.519	0.760	0.903	0.968	0.991	0.998
-0.4	0.000	0.000	0.012	0.085	0.267	0.515	0.738	0.884	0.957	0.987
-0.3	0.000	0.000	0.002	0.025	0.110	0.284	0.511	0.722	0.869	0.949
-0.2	0.000	0.000	0.000	0.006	0.037	0.128	0.294	0.508	0.711	0.860
-0.1	0.000	0.000	0.000	0.001	0.010	0.047	0.139	0.299	0.504	0.704
0.0	0.000	0.000	0.000	0.000	0.002	0.014	0.054	0.145	0.300	0.500
0.1	0.000	0.000	0.000	0.000	0.000	0.003	0.017	0.056	0.145	0.296
0.2	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.017	0.055	0.140
0.3	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.016	0.051
0.4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.013
0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002

Table A.1: CDF table for correlation coefficient, left-tail

$\rho\backslash$ r	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
-0.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.5	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	0.987	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.3	0.949	0.984	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000
-0.2	0.860	0.945	0.983	0.996	0.999	1.000	1.000	1.000	1.000	1.000
-0.1	0.704	0.855	0.944	0.983	0.997	1.000	1.000	1.000	1.000	1.000
0.0	0.500	0.700	0.855	0.946	0.986	0.998	1.000	1.000	1.000	1.000
0.1	0.296	0.496	0.701	0.861	0.953	0.990	0.999	1.000	1.000	1.000
0.2	0.140	0.289	0.492	0.706	0.872	0.963	0.994	1.000	1.000	1.000
0.3	0.051	0.131	0.278	0.489	0.716	0.890	0.975	0.998	1.000	1.000
0.4	0.013	0.043	0.116	0.262	0.485	0.733	0.915	0.988	1.000	1.000
0.5	0.002	0.009	0.032	0.097	0.240	0.481	0.760	0.946	0.997	1.000
0.6	0.000	0.001	0.005	0.020	0.071	0.208	0.477	0.804	0.980	1.000
0.7	0.000	0.000	0.000	0.002	0.009	0.042	0.163	0.474	0.874	0.999
0.8	0.000	0.000	0.000	0.000	0.000	0.002	0.014	0.098	0.470	0.969
0.9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.021	0.466

Table A.2: CDF Table for correlation coefficient, right-tail

Appendix B

Matlab Codes

B.1 rcdf.m

```matlab
function p = rcdf(T,rho,z)
    % \rho \in [-1,1]
    if z > 1 || z < -1 error('[!Correlation is between -1 and 1!]'); end
        % Probability Density Function
        function p0 = rpdf(T,rho,r)
        p0 = (T-2)*gamma(T-1)*(1-rho^2)^((T-1)/2)*(1-r.^2)^((T-4)/2)/ ...
        (sqrt(2*pi)*gamma(T-1/2)*(1-rho.*r)^(T-3/2))* ...
        hypegeometric(1/2,1/2,(2*T-1)/2,(rho*r+1)/2);
    end
    % Riemann Sum to approximate the CDF
    dx = 1e-4; r = -1:dx:z;
    for i = 1: length(r) R(i) = rpdf(T,rho,r(i)); end
    p = trapz(r,R);
end
```

B.2   hypergeometric.m

```
function F = hypergeometric(a,b,c,x,tol)
    % If tolerance is not specified, use 1e-7 as default
    if nargin == 4 tol = 1e-10;
    elseif nargin ~= 5 error('Wrong Number of Arguments');
    end
    s0 = 0; s1 = 1; j = 1;
    while abs(s1-s0) > tol
        s0 = s1;
        s1 = s1 + gamma(a+j)/gamma(a)*gamma(b+j)/gamma(b)* ...
        gamma(c)/gamma(c+j)*x^j/factorial(j);
        j = j + 1;
    end
    F = s1;
end
```

B.3   wishlow.m

```
function Z = wishlow(Sigma, df, alpha)
    n = length(Sigma); [T D1] = wishrnd(Sigma, df);
    X = chi2inv(1-alpha,n*(n+1)/2);
    while (1)
        Z = wishrnd(Sigma,df,D1)/df;
        L = df*(log(det(Z))-log(det(Sigma))+trace(Sigma*inv(Z))-n);
        L = (1-1/(6*df)*(2*n+1-2/(n+1)))*L;
        if L <= X && abs(L-X) < 10e-5 break; end
    end
end
```

References

References

[1] Theodore Anderson, "An Introduction to Multivariate Statistical Analysis", *John Wiley & Sons Inc.*, 1985

[2] Maurice Bartlett, "A Note on Multiplying Factors for Various Chi-squared Approximations", *Journal of Royal Statistical Society, Series B*, Vol. 16, pp. 296-298, 1954

[3] Richard Brealey and Stewart Hodges," Portfolio Selection in a Dynamic and Uncertain World", *Financial Analysts Journal*, Vol. 28, No. 6, pp. 58-69, 1972

[4] Zengjing Chen and Larry Epstein, "Ambiguity, Risk, and Asset Returns in Continuous Time", *Econometrica*, Vol. 70, No. 4, pp. 1403-1443, 2002

[5] Daniel Ellsberg, "Risk, Ambiguity, and the Savage Axioms", *The Quarterly Journal of Economics*, Vol. 75, No. 4, pp. 643-669, 1961

[6] Larry Epstein, "A Definition of Uncertainty Aversion", *The Review of Economics Studies*, Vol. 66, pp. 579-608, 1999

[7] Frank Fabozzi, "Robust Portfolio Optimization and Management", *John Wiley & Sons, Inc.*, 2007

[8] Lorenzo Garlappi, Raman Uppal and Tan Wang, "Portfolio Selection with Parameter and Model Uncertainty: A Multi-prior Approach", *The Review of Financial Studies*, Vol. 20, No. 1, pp. 41-81, 2007

[9] Raymond Kan and Guofu Zhou, "Optimal Portfolio Choice with Parameter Uncertainty", *Journal of Financial and Quantitative Analysis*, Vol. 42, No. 3, pp. 621-656, 2007

[10] Yoram Kroll and Haim Levy, "Sampling Errors and Portfolio Efficient Analysis", *Journal of Financial and Quantitative Analysis*, Vol. 15, Issue 03, pp. 655-688, 1980

[11] Hiroshi Konno and Hiroaki Yamazaki, "Mean-Absolute Deviation Portfolio Optimization Model and Its applications to Tokyo Stock Market", *Management Science*, Vol. 26, No. 11, pp. 1108-1116, 1980

[12] Hening Liu, "Dynamic Portfolio Choice Under Ambiguity and Regime Switching Mean Returns", *Journal of Economic Dynamic & Control*, Vol. 35, pp. 623-640, 2010

[13] Harry Markowitz, "Portfolio Selection, *The Journal of Finance*", Vol. 7, No. 1, pp. 77-91, 1952

[14] Richard Michaud, "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?", *Financial Analysts Journal*, 1989

[15] Donald Morrison, "Multivariate Statistical methods", *McGraw-Hill Book Company*, 2nd edition, 1967

[16] Anatolii Puhalskii, "On Portfolio Choice by Maximizing the Outperformance Probability", *Mathematical Finance*, Vol. 21, No. 1, pp. 145-167, 2011

[17] Marc Oliver Rieger and Mei Wang, "Cumulative Prospect Theory and the St. Petersburg Paradox", *Economic Theory*, Vol. 28, No. 3, pp. 665-679, 2006

[18] Yusif Simaan, "Estimation Risk in Portfolio Selection: The Mean Variance Model Versus the Mean Absolute Deviation Model", *Management Science*, Vol. 43, No. 10, pp. 1437-1446, 1997