T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Ex-post realized Sharpe ratio and information ratio achieved by the ex-ante forecast optimal portfolio
著者(和文)	清水正人
Author(English)	Masahito Shimizu
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11410号, 授与年月日:2020年3月26日, 学位の種別:課程博士, 審査員:中野 張,二宮 祥一,尾形 わかは,仙石 慎太郎,中丸 麻由子,中川 秀敏
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11410号, Conferred date:2020/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	要約
Type(English)	Outline

Powered by T2R2 (Science Tokyo Research Repository)

Ex-post realized Sharpe ratio and information ratio achieved by the ex-ante forecast optimal portfolio

Masahito Shimizu
Graduate School of Innovation Management, Tokyo Institute of Technology

Abstract

The ex-post realized Sharpe ratio and the information ratio using an ex-ante forecast return distribution—possibly different from the ex-post realized distribution—are studied in this thesis. Forecasting power measurements are proposed, as are attribution formulas and methods. The proposed methods enable the ex-post realized Sharpe ratio and information ratio to be decomposed into realized market conditions, the ex-ante predictability of returns, the risk magnitude, and the risk factors. The methods enable a fair and direct comparison of the proposed measurement of the ex-ante return and ex-ante risk, quantitative identification of the main source of the reduction in the Sharpe ratio from the achievable Sharpe ratio, and a reduction in the information ratio from the achievable information ratio. The proposed method performs an excess Sharpe ratio attribution analysis that simultaneously evaluates the qualities of a portfolio and the benchmark. Numerical examples of the attributions are provided using sector indices. The examples show that the ex-ante risk forecast has a different effect between the predictability of the Sharpe ratio and the predictability of the information ratio. As an application of the provided measurement, an implication in the optimal confidence level in the Black-Litterman model is provided.

Contents

Chapter 1

Introduction

1.1 Sharpe ratio and information ratio in modern portfolio theory

Making an investment is a decision-making process. A portfolio that has investable assets with uncertain performance is selected based on an investor's forecasts at the beginning of an investment period. The portfolio is exposed to various economic and market conditions. The portfolio's achieved performance refers to the investment results during and at the end of the investment period, and the performance is evaluated. In this thesis, relations between forecasts at that moment of portfolio selection and measurements that evaluate performance are studied, and the causes of the performance are studied. Forecasts of the returns and risks of the investable assets are set as the forecast, and measurements of the return per unit of risk are set as the measurements. The Sharpe ratio and the information ratio are studied as the measurements.

Markowitz (1952) states that "the investor does (or should) consider expected return a desirable thing *and* variance of return an undesirable thing" in the process of selecting a portfolio. Thus, expected returns and variance of returns are important concepts in modern portfolio theory, and an optimal portfolio is constructed from the investor's expected returns and risk. In modern portfolio theory, the variance of returns is recognized as a measure of risk.

Performance evaluations are important activities for portfolio managers and investors regarding investments. Elton and Gruber (1987) state that "an integral part of any decision-making process should be the evaluation of the decision. This is equally true whether investors make their own investment decisions or employ a manager to make them." Farrell (1983) states, "Performance evaluation aims not only at assessing the success of the investment process in achieving the overall investment goals but also at diagnosing the contribution of the individual elements that have made possible achievement of the overall goal." Grinold and Kahn (1999) similarly state, "a sophisticated performance analysis system can provide valuable feedback for an active manager." Performance evaluations are important activities for future investments, as previously mentioned. Furthermore, the importance of analyzing the contribution of the individual elements of investment decisions is widely acknowledged. Performance attribution analyses that identify strengths and weaknesses of the elements are performed in many aspects of investment analysis.

The Sharpe ratio is an important measure of the return per unit of risk. Lo (2002) states, "One of the most commonly cited statistics in financial analysis is the Sharpe ratio, the ratio of the excess expected return of an investment to its return volatility or standard deviation. Originally motivated by mean-variance analysis and the Sharpe-Lintner Capital Asset Pricing Model, the Sharpe ratio is now used in many different contexts, from performance attribution to test of market efficiency to risk management." Chan (2017) similarly states, "The Sharpe ratio and the maximum drawdown are two of the most important tools to measure the risk of an investment. Since Sharpe introduces the Sharpe ratio, which is the excess return per unit of volatility, it is widely used to evaluate the performance of portfolios."

The Sharpe ratio is defined as the quantity of excess return per unit quantity of risk.

$$\text{Sharpe ratio} \equiv \frac{\text{the excess return}}{\text{the standard deviation of the excess return}}$$

Excess return is defined as the return of a portfolio minus the return of a risk-free asset. Additionally, risk is measured by the standard deviation of the excess return.

In performance analyses, a portfolio's performance is often compared with a benchmark performance. Especially, a portfolio which declares its benchmark is always compared with the benchmark. Additionally, there are investments called as enhanced index investments. These investments propose to achieve a return of its benchmark's one plus more positive return. In such cases, a portfolio's performance is evaluated using the information ratio. The information ratio is defined as the quantity of a portfolio's excess return against the return of a benchmark per unit quantity of risk. Additionally, risk is measured by the standard deviation of the excess return. Thus, the information ratio is defined as the quantity of excess return per unit quantity of risk.

$$\text{Information ratio} \equiv \frac{\text{the excess return against the benchmark return}}{\text{the standard deviation of the excess return against the benchmark return}}$$

Regarding the expanded uses of the information ratio, Grinold and Kahn (1999) state, "The institutional investor in the United States and to an increasing extent worldwide cares about active risk and active return. For that reason, we will concentrate on the more general problem of managing relative to a benchmark." They continue to state as follows. "In fact, analyzing investments relative to a benchmark is more general than the standard total risk and return framework."

Additionally, Sharpe ratio and the information ratio are widely accepted as important risk-adjusted performance measures in fields of investment business practices.

Studies on the information ratio follow the study in this thesis of the Sharpe ratio.

1.2 Organization of this thesis

A portfolio that has investable assets with uncertain performance is selected based on an investor's forecasts at the beginning of an investment period, and the portfolio's achieved performance refers to the investment results during and at the end of the investment period, and the performance is evaluated. The investor's forecasts at the beginning of the period are ex-ante forecasts, and it is natural that the forecasts are not necessarily identical to the performance of investable assets during the period. Therefore, the portfolio's achieved performance is brought from the performance of investable assets during the period. The performances of the assets are ex-post realized returns which are observed at the end of the period. This thesis studies the relations between ex-ante forecasts of the future performance of investable assets at the beginning of an investment period and measurements of the portfolio's ex-post performance analysis. Ex-ante forecast returns and risk of investable assets are set as the ex-ante forecasts, and the ex-post realized Sharpe ratio and the information ratio are set as measurements of the ex-post performance analysis.

The causes and sources of the performance and analysis of risk and return are important issues for performance evaluation. Studies on attribution analyses in the performance analyses of portfolios are another important theme of this thesis.

Related previous works are reviewed in Chapter 2. Themes on portfolio construction and performance analyses are often independently studied in textbooks, such as those written by Farrell (1983) and Elton and Gruber (1987). Nevertheless, Steiner (2011) studies a Sharpe ratio attribution analysis, which represents research on the relation between the elements of a portfolio at that moment of its construction and ex-post Sharpe ratio attribution. Steiner decomposes the Sharpe ratio into risk weights, diversification effects, and asset Sharpe ratios—elements of the portfolio's properties at the beginning of the investment period.

In this thesis, ex-ante forecast returns and risk are set as the elements of a portfolio at that moment of its construction. As Markowitz (1952) states, "the ex-ante forecast returns and risk provide the weights of the investable assets in the portfolio." The weights of the investable assets provide the risk weights and diversification effects described in Steiner (2011). Thus, ex-ante forecast returns and risk as the sources of the weights are objects of study in Chapter 3 and Chapter 4.

Grinold and Kahn (1999) study the cause of an information ratio. They show that the coefficient between signals that predict returns of investable assets in a future investment period and returns realized in the investment is one source of the cause of the information ratio. In Grinold and Kahn (1999), this property is called the fundamental law of active management.

Menchero (2016) studies the cause of reducing an information ratio. He investigates effects brought by differences of information contained in forecast returns and risk which an investor uses in constructing the optimal portfolio using the mean-variance approach. He shows the effects by empirical exercise using historical data samples. Chan *et al*. (2017) show an upper bound for ex-post Sharpe ratio. The result is an inequality which is a formula of parameters about time series of returns given, which include the maximum drawdown of the returns. Additionally, they show numerical analysis using historical data. However, these studies do not focus on investor's forecasts for construction of optimal portfolios.

The formulation of this study is provided in Chapter 3 and is composed of two steps. The first step is summarized in Theorem 3.1. If an investor constructs an optimal portfolio that achieves the expected maximum Sharpe ratio based on ex-ante forecast returns and risk, Theorem 3.1 indicates the returns that provide the identical portfolio to the optimal portfolio under ex-post risk, for which the investor constructs the optimal portfolio by using the mean-variance approach of modern portfolio theory. The second step is calculating the Sharpe ratio achieved by the optimal portfolio under ex-post realized returns and risk. In the second step, a measurement that evaluates the power of an ex-ante forecast in predicting future returns and risk is proposed. In this thesis, we call the measurement *the predictability-based measurement*, which is a number between one and negative one. The highest Sharpe ratio is achieved when this number is one, and the lowest Sharpe ratio when this number is negative one. Additionally, we find that the Sharpe ratio is the product of the proposed predictability-based measurement multiplied by the absolute value of the market Sharpe ratio. The absolute value of the market Sharpe ratio during the investment period is recognized as market conditions of the period. The first contribution in Chapter 3 is to formulate Theorem 3.1, which connects an optimal portfolio based on ex-ante forecast returns and risk and an optimal portfolio based on ex-post realized returns and risk. The second contribution is to introduce a predictability-based measurement to simultaneously evaluate the ex-ante forecast risk and return. The third contribution is to propose the predictability-based measurement, which is independent of market conditions.

Numerical examples are provided in Chapter 3. S&P sector total return indices are used for numerical examples. S&P indices are recognized as one of the widely used indices for evaluating investment markets, equity markets, fixed income markets, and so on. S&P Dow Jones Indices (2018) states, "Our sector and industry indices measure the performance of the widely-used Global Industry Classification Standard (GICS®) sectors and sub-industries. GICS enables market participants to identify and analyze a customized group of companies using a common global standard." Additionally, many index-type investment vehicles, index funds, exchange-traded funds, and investment accounts exist and with performances that mimic index performances. Thus, the sector total return indices are almost investable. The proposed predictability-based measurements of the three sample forecast returns are evaluated for five years. In the examples, we find the importance of independence from market conditions for the proposed predictability-based measurement.

Investments during one period are discussed from Chapter 3.1 to 3.4. The discussions are extended for multi-period investments in Chapter 3.5. Under a certain assumption, similar discussions are completed for multi-period investments. Analysis described in Chapter 3 has been published in Shimizu, M. (2017).

A Sharpe ratio attribution to the ex-ante forecast returns and risk is studied in Chapter 4. First, the ex-post realized Sharpe ratio is attributed to the ex-ante forecast returns and risk. Next, the ex-ante forecast risk is attributed to risk factors and the risk magnitudes of the risk factors.

Risk factors are often studied in modern portfolio theory. Fama and French (1993) study five common risk factors in stock and bond returns. Three stock-market factors exist: an overall market factor and factors related to firm size and book-to-market equity. They state that stock returns have shared variations given the stock-market factors. Ross (1976) proposes arbitrage pricing theory, which is based on a multi-factor model. The Barra risk model is a widely used risk model in investment practices. Regarding a risk model, MSCI (2018) states, "We need measures of security dispersion and co-movement. These are contained in the covariance matrix of security returns. There are a number of ways of estimating the covariance matrix of security returns. Substantial gains are made by recognizing that covariance is driven by common sources of returns across securities. These common sources of returns are called common factors. Estimating the covariance matrix of security returns thus depends on estimating a factor model for security returns."

Therefore, the attribution of ex-ante forecast risk decomposes into the attribution of risk factors and of the risk magnitude of the risk factors. The first contribution in Chapter 4 is to provide an ex-post realized Sharpe ratio attribution to ex-ante forecast returns and risk. The second contribution is to provide the attribution using a formula. Furthermore, we decompose the predictability-based measurement of the risk into the predictability-based measurement of the risk factors and the predictability-based measurement of the magnitudes of the factors. This attribution formula enables us to quantitatively compare the predictability-based measurement of the return with the predictability-based measurement of the risk. The third contribution is to provide the excess Sharpe ratio attribution against a given benchmark using the predictability-based measurement of the implied returns forecast of the benchmark portfolio. Hence, we find the relative strength of an investor's portfolio against a benchmark and the absolute strengths of the portfolio and the benchmark, respectively. Consequently, the proposed attribution analysis derives rich information. Numerical examples are also provided in Chapter 4. Analysis described in Chapter 4 is published in Shimizu, M. (Forthcoming). An outline and results of this chapter of the thesis is described because of including the contents published in the future in Shimizu, M. (Forthcoming).

In Chapter 5, an application of the proposed predictability-based measurement for the Black-Litterman model is provided.

The Black-Litterman model is introduced because problems are encountered in the mean-variance optimization in practice. Regarding mean-variance optimization in practice, Fabozzi et al. (2006) state, "Consequently, these 'optimized' portfolios are not necessarily well diversified and exposed to unnecessary *ex post* risk. The reason for these phenomena is not a sign that mean-variance optimization does not work, but rather that the modern portfolio theory framework is very sensitive to small changes in the input," and, "In practice, the Black-Litterman model has proven successful in mitigating estimation errors in the mean-variance framework. The Black-Litterman model does not require that the portfolio manager provide forecasts on all the securities in the universe managed. Rather, it 'blends' any views (this could be a forecast on just one or a few securities, or all them) the investor might have with market equilibrium."

The Black-Litterman model has a key parameter, which is the confidence level of the investor's view. The confidence level of the investor's view indicates the relative importance of equilibrium versus the views. The confidence level varies the proposed predictability-based measurement of expected returns brought by the Black-Litterman model. In Chapter 5, a numerical example of this phenomenon is provided in which strong investor confidence reduces the realized Sharpe ratio and not the maximum Sharpe ratio. The phenomenon in the numerical example indicates an issue when considering the confidence level of the Black-Litterman model.

In Chapter 6, the relation between ex-ante forecast returns and risk and the ex-post realized information ratio is studied. A measurement that evaluates the power of an ex-ante forecast that predicts the ex-post realized information ratio is proposed, similar to the case of the Sharpe ratio. In this thesis, we call the measurement the predictability-based measurement for the information ratio and is a number between one and negative one. An attribution of the ex-post realized information ratio is studied, similar to discussions in Chapter 3 and Chapter 4. Active portfolios are discussed in Chapter 6, where they are defined as the excess weight vectors against the weight vector of a benchmark portfolio. First, Theorem 6.1 is provided, which shows the returns for the optimal portfolio that achieves an information ratio under an ex-post risk identical to the information ratio of the optimal portfolio that an investor constructs under ex-ante returns and risk using the mean-variance approach. Theorem 6.1 also shows that the predictability-based measurement of the perfect ex-ante forecast, with forecast returns and risk identical to the ex-post realized returns and risk, is not necessarily equal to one. The ex-post realized information ratio is attributed to the ex-ante forecast returns and risk, similar to Chapter 4. Numerical examples are also provided in Chapter 6. The first contribution in Chapter 6 is to formulate Theorem 6.1. The second contribution is to provide an ex-post realized information ratio attribution to the ex-ante forecast returns and risk. Furthermore, we decompose the predictability-measurement of risk into risk factors and the risk magnitude of the risk factors. The third contribution in Chapter 6 is to show that the predictability-based measurement of the perfect forecast is not necessarily equal to one. This result indicates that applying other performance measures such as the Sharpe ratio in addition to the information ratio to evaluate the predictive power of a forecast is desirable.

In Chapter 7, we summarize the studies and find that future studies are needed. The Appendices that provide the derivations are included at the end of this thesis.

Chapter 2

Related previous works

2.1 Sharpe ratio attribution analysis

Steiner (2011) points out that "classical performance attribution approaches, e.g. Brinson decompositions, ignore risk aspect altogether." Steiner proposes a Sharpe ratio contribution analysis that has "information on portfolio constituents available to analysts" and considers risk aspects.

In this section, we review the Sharpe ratio attribution analysis in accordance with Steiner (2011). The Sharpe ratio S_p of portfolio P is defined as the quotient of portfolio excess return r_p and portfolio volatility σ_p

$$S_p \equiv \frac{r_p}{\sigma_p}. \tag{2.1.1}$$

r_p is decomposed linearly into asset contributions and is calculated as asset weights w_i multiplied by asset excess return r_i for the $i-\text{th}$ asset in the portfolio. Therefore, we have

$$S_p = \sum_i \frac{w_i r_i}{\sigma_p}. \tag{2.1.2}$$

Because volatility is a linear homogenous function, the weighted marginal contributions to volatility sum up to portfolio volatility:

$$\sigma_p = \sum_i w_i \frac{\partial \sigma_p}{\partial w_i}. \tag{2.1.3}$$

The marginal contribution of an asset to portfolio volatility is its covariance $\sigma_{i,p}$ with the overall portfolio. From the definition of covariance, it follows that the marginal contribution of an asset to portfolio risk can be expressed as

$$\frac{\partial \sigma_p}{\partial w_i} = \frac{\sigma_{i,p}}{\sigma_p} = \rho_{i,p}\sigma_i, \tag{2.1.4}$$

where $\rho_{i,p}$ is the correlation coefficient. This expression can be summarized with the following portfolio volatility decomposition

$$\sigma_p = \sum_i w_i \rho_{i,p} \sigma_i. \tag{2.1.5}$$

From (2.1.5), we have

$$S_p = \sum_i \frac{w_i r_i}{\sigma_p} \frac{w_i \rho_{i,p} \sigma_i}{\sigma_p} \frac{\sigma_p}{w_i \rho_{i,p} \sigma_i} = \sum_i \frac{w_i \rho_{i,p} \sigma_i}{\sigma_p} \frac{1}{\rho_{i,p}} \frac{r_i}{\sigma_i}. \tag{2.1.6}$$

Equation (2.1.6) is an additive decomposition of the Sharpe ratio, in which an asset's contribution consists of three components:

$$\text{Risk weights } \frac{w_i \rho_{i,p} \sigma_i}{\sigma_p}. \tag{2.1.7}$$

$$\text{Diversification effects } \frac{1}{\rho_{i,p}}. \tag{2.1.8}$$

$$\text{Asset Sharpe ratios } \frac{r_i}{\sigma_i}. \tag{2.1.9}$$

Equation (2.1.7) are weights because they sum to unity:

$$\sum_i \frac{w_i \rho_{i,p} \sigma_i}{\sigma_p} = 1. \tag{2.1.10}$$

Equation (2.1.8) reflects the interaction of the assets when added to the portfolio.

Steiner decomposes the Sharpe ratio into risk weights, diversification effects, and asset Sharpe ratios.

2.2 The fundamental law of active management

Grinold and Kahn (1999) discuss the information ratio and provide a formula that expresses the ratio using two parameters. One parameter is the correlation between the returns of risk assets and the values that forecast the returns. In Grinold and Kahn (1999), the values that forecast the returns are called signals. The correlation is called an *information coefficient* (IC). The other parameter is the number of signals, called breadth in Grinold and Kahn (1999). They show that the information ratio is expressed as the product of the IC and the breadth of the square root of the strategy. Grinold and Kahn (1999) provide a formula called the fundamental law of active management. In their discussion of the IC, the returns are residual and specific returns that do not result from systematic risks, such as market risk. We note that the IC is not necessarily the coefficient of residual returns. In Farrell (1983), we find an IC between the actual attractiveness of the rankings and the forecast attractiveness of the rankings for investable assets.

In this section, we provide a brief review of the fundamental law of active management in accordance with Grinold and Kahn (1999). Detail derivations are given in Appendix 2.4.

Suppose that the residual return $\boldsymbol{\theta}$ for n risky assets is described by

$$\boldsymbol{\theta} = \boldsymbol{A}\boldsymbol{x} \tag{2.2.1}$$

where

$$\boldsymbol{x} \equiv \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix}, \tag{2.2.2}$$

and $\boldsymbol{x}_i$ represents random variables under a normal distribution $\boldsymbol{N}(0,1)$. Further, each pair $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ is assumed to be independent.

Suppose that an investor has BR signals $\boldsymbol{z} \equiv \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_{\mathrm{BR}} \end{pmatrix}$ for the residual return $\boldsymbol{\theta}$ and that

$$\boldsymbol{z} = \boldsymbol{E}\boldsymbol{y} \tag{2.2.3}$$

where

$$\boldsymbol{y} \equiv \begin{pmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{\mathrm{BR}} \end{pmatrix}, \tag{2.2.4}$$

and $\boldsymbol{y}_i$ represents the BR random variables under $\boldsymbol{N}(0,1)$. Further, each pair $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ is assumed to be independent.

We consider that the signals $\boldsymbol{z}$ have information for the residual returns $\boldsymbol{\theta}$, which indicates that we have the regression analysis result of $\theta_i = a + bz_i + \varepsilon_i$. Let $\rho_{n,b}\ n = 1, \cdots, n, b = 1, \cdots, \mathrm{BR}$ be the correlation of each signal element $\boldsymbol{y}_j$ with the basis element $\boldsymbol{x}_i$. Using the approximation formula for a matrix, we define the sum of the correlation of each signal with the basis element over the basis element:

$$\zeta_b^2 \equiv \sum_{n=1}^{N} \rho_{n,b}^2. \tag{2.2.4}$$

By assuming that all the signals have equal value

$$\zeta_b^2 = \rho^2 = \mathrm{IC}^2, \tag{2.2.5}$$

Grinold and Kahn (1999) provide the fundamental law of active management:

$$\mathrm{IR}^2 = \mathrm{BR} \cdot \mathrm{IC}^2 \tag{2.2.6}$$

The derivation uses a matrix inversion approximation and the assumption that all signals have equal value.

Chapter 3

A predictability-based measurement for the Sharpe ratio

3.1 Introduction

Improving the Sharpe ratio is an important goal of sponsors and portfolio managers. Investors manage their portfolios by relying on their ex-ante forecast return distributions for investable assets at the beginning of an investment period. As such, it is natural that the ex-ante forecast return distribution differs from the ex-post return distributions or the realized future distributions.

Consider a one-time investment in n risky assets, where expected returns are distributed according to an n-dimensional normal distribution. The purpose of the investment is to maximize the Sharpe ratio using the mean-variance approach. After the investment period, we evaluate the ex-ante forecast.

Grinold and Kahn (1999) introduce the *information coefficient* (IC) as a measure of forecasts. They define the IC as the correlation of each forecast with its actual outcome. If the IC is the same for all forecasts, Grinold and Kahn (1999) use the *fundamental law* *of active management* to show that the information ratio is the product of the IC and the breadth of the square root of the strategy. Here, the breadth is defined as the number of independent forecasts that generate the exceptional returns that an investor makes per year. They then analyze the relationship between a risk-adjusted performance measurement and measurements that appear in the investment process.

Originally, the IC is the correlation between the residual return of each forecast and its actual outcome. However, we also evaluate how closely we forecast the ex-post return of each investable asset. Furthermore, only a close ex-ante forecast return cannot achieve a high Sharpe ratio. We also need a risk forecast that uses the mean-variance approach, which plays an important role in portfolio construction processes.

For example, we assume that we have three investable assets: Asset A, Asset B, and Asset C. We assume that Asset A has an ex-post return of $r_A = 6\%$, Asset B of $r_B = 3\%$, and Asset C of $r_C = 1\%$, respectively. Regarding their ex-post risks of returns, the three assets show the ex-post standard deviations $\sigma_A = 8\%, \sigma_B = 8\%,$ and $\sigma_C = 2\%$, the returns of Asset A and Asset B have the correlation $\rho_{AB} = 0.6$, the returns of Asset A and Asset C are independent, and the returns of Asset B and Asset C are also independent. Given these return distributions, we can construct the optimal portfolio that is 31% long on Asset A, 4% short on Asset B, and 73% long on Asset C. The optimal portfolio has a maximum Sharpe ratio of 0.91.

Suppose an investor has a perfect ex-ante return forecast and an imperfect ex-ante risk forecast. We assume that the investor forecasts that Asset A has an ex-ante return of $r_A = 6\%$, Asset B of $r_B = 3\%$, and Asset C of $r_C = 1\%$, respectively. Additionally, we assume that the investor forecasts the ex-ante standard deviations $\sigma_A = 10\%, \sigma_B = 5\%,$ and $\sigma_C = 2\%$ and that each pair of returns is independent. However, the investor must construct the optimal portfolio from his or her ex-ante forecast return distribution using the mean-variance approach. The investor builds the optimal portfolio with 14% long on Asset A, 28% long on Asset B, and 58% long on Asset C positions. The optimal portfolio achieves a Sharpe ratio of 0.70 under the ex-post realized return distribution. The optimal portfolio delivers a Sharpe ratio that is approximately 75% of the maximum Sharpe ratio of the optimal portfolio from the ex-post realized distribution. This simple example shows that it is not enough to only evaluate the ex-ante forecast return.

We consider a measure that evaluates the ex-ante forecast with respect to risk and return. We first show that the optimal portfolio based on an ex-ante forecast return distribution is identical to the optimal portfolio based on the "risk basis transposed forecast return" and the ex-post realized covariance matrix; the "risk basis transposed forecast return" is introduced in the second section. Theorem 3.1 summarizes this fact.

Note that the "risk basis transposed forecast return" includes information on the ex-ante forecast return as well as the difference between the ex-ante and ex-post covariance matrices. Next, we calculate the Sharpe ratio for an optimal portfolio based on the return and the achieved ex-post realized covariance matrix. We then introduce a measure to evaluate the return. Applying the "risk basis transposed forecast return" to the measure, we obtain a measure that evaluates the ex-ante forecast with respect to risk and return. The measure is a number between one and negative one. The highest Sharpe ratio is achieved when this number is one, and the lowest Sharpe ratio when this number is negative one. Additionally, as the number increases, the Sharpe ratio of the portfolio increases. Finally, we show that the Sharpe ratio is the product of the measure multiplied by the absolute value of the Sharpe ratio of the market, indicating that the measure is separate from market conditions.

The first contribution of this chapter is to propose a measure to evaluate simultaneously the ex-ante forecast risk and return. The second contribution is to achieve a better understanding of the Sharpe ratio and show that this ratio is the product of the introduced measure multiplied by the absolute value of the market Sharpe ratio. Finally, we highlight the importance of both the risk and return forecasts.

This chapter is organized as follows. In Section 3.2, we introduce the formulas and Theorem 3.1. In the third section, we introduce a measure to evaluate an ex-ante forecast distribution, which is calculated at the end of the investment period, and describe the properties of the measure. Then, we provide numerical examples of the measure using an investment in ten weekly S&P sector total return indices by evaluating three ex-ante forecasts that are implied from the market weight portfolio, the equal weight portfolio, and the minimum variance portfolio. Furthermore, we recognize the importance of evaluating the forecast separate from market conditions that investors cannot consider at the time of portfolio construction. Finally, we discuss the implications for practical investment actions and the need for further research. The derivations are provided at the end of this paper.

3.2 Formulation

We consider one-period investment in a non-leveraged long/short portfolio. We have n risky assets and construct the portfolio using the mean-variance approach without leverage to maximize the expected Sharpe ratio. For simplification, we assume a constituent with a small investment in the portfolio and trade without friction. For this paper, "return" refers to the excess return to the risk-free return. We also suppose that the rank of the covariance matrix of both the ex-ante and ex-post n-dimensional normal distributions is n. Therefore, both distributions have n orthonormal risk factors.

Suppose that the n risky assets behave according to an n-dimensional normal distribution $\boldsymbol{N(r,\Sigma)}$. First, we must find the optimal portfolio to maximize the expected Sharpe ratio under the distribution $\boldsymbol{N(r,\Sigma)}$.

Let $\boldsymbol{w} = (w_1, \cdots, w_n)'$ be a vector of the portfolio weights of risky assets, where $\boldsymbol{x}'$ denotes the transposed vector of $\boldsymbol{x}$. For a given number h, we find the minimum-variance portfolio as the solution to the following quadratic programming problem:

$$minimize\ \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w},$$

$$\text{s.t. } \boldsymbol{r}'\boldsymbol{w} = h, \text{and } \boldsymbol{e}'\boldsymbol{w} = 1, \tag{3.2.1}$$

where $\boldsymbol{e} \equiv (1, \cdots, 1)'$. Suppose that the solution to (3.2.1) is $\boldsymbol{w}_h$. The Sharpe ratio of the portfolio is $f(h) \equiv h/\sqrt{\boldsymbol{w}_h{}'\boldsymbol{\Sigma}\boldsymbol{w}_h}$ and is a function of h. Then, we find h_0, which provides the maximum value of $f(h)$ under the condition that $\boldsymbol{e}'\boldsymbol{\Sigma}^{-1}\boldsymbol{r}$ is positive. Let $\boldsymbol{w}_0$ be $\boldsymbol{w}$ when $h = h_0$. Indeed, $\boldsymbol{w}_0$ is the weight vector of the optimal portfolio that achieves the maximum Sharpe ratio. In that case, we obtain (see Appendix 1):

$$\boldsymbol{w}_0 = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{r}}{\boldsymbol{e}'\boldsymbol{\Sigma}^{-1}\boldsymbol{r}}. \tag{3.2.2}$$

Suppose we have an ex-ante forecast on n risky assets in the future investment period. Therefore, we have an ex-ante forecast with an n-dimensional normal distribution $\boldsymbol{N}(\boldsymbol{r_F},\boldsymbol{\Sigma_F})$ at the beginning of the investment period. Although the n risky assets behave according to an ex-post realized n-dimensional normal distribution $\boldsymbol{N}(\boldsymbol{r_R},\boldsymbol{\Sigma_R})$ during the period, we do not have perfect information about the future. That $\boldsymbol{r_F}$ is not identical to $\boldsymbol{r_R}$ and $\boldsymbol{\Sigma_F}$ is not identical to $\boldsymbol{\Sigma_R}$ is natural. From (3.2.2), we have the optimal portfolio $\boldsymbol{w_F}$ based on the ex-ante forecast distribution $\boldsymbol{N}(\boldsymbol{r_F},\boldsymbol{\Sigma_F})$

$$\boldsymbol{w}_F = \frac{\boldsymbol{\Sigma_F^{-1}r_F}}{\boldsymbol{e}'\boldsymbol{\Sigma_F^{-1}r_F}}. \tag{3.2.3}$$

Recall the assumption for the rank of the two-covariance matrix. We can find the orthonormal matrixes $\boldsymbol{K_F}$, and $\boldsymbol{K_R}$ such that they diagonalize $\boldsymbol{\Sigma_F}$, and $\boldsymbol{\Sigma_R}$, respectively. Then, we obtain

$$\boldsymbol{K_F^{-1}\Sigma_F K_F} = \boldsymbol{\sigma_F}^2, \boldsymbol{K_R^{-1}\Sigma_R K_R} = \boldsymbol{\sigma_R}^2, \tag{3.2.4}$$

where

$$\boldsymbol{\sigma_F} = \begin{pmatrix} \sigma_F(1) & & 0 \\ & \ddots & \\ 0 & & \sigma_F(n) \end{pmatrix}, \boldsymbol{\sigma_R} = \begin{pmatrix} \sigma_R(1) & & 0 \\ & \ddots & \\ 0 & & \sigma_R(n) \end{pmatrix}. \tag{3.2.5}$$

Indeed, $\boldsymbol{K_F} = \left(\boldsymbol{c_F}(1), \cdots, \boldsymbol{c_F}(n)\right)$ is a matrix that contains the orthonormal ex-ante forecast factor risk basis $\boldsymbol{c_F}(1), \cdots, \boldsymbol{c_F}(n)$ as the column vectors, and $\boldsymbol{K_R} = (\boldsymbol{c_R}(1), \cdots, \boldsymbol{c_R}(n))$ is a matrix that contains the orthonormal ex-post realized risk factor basis $\boldsymbol{c_R}(1), \cdots, \boldsymbol{c_R}(n)$ as the column vectors.

We introduce a "risk basis transposer" $\boldsymbol{\Phi}$ as follows:

$$\boldsymbol{\Phi} \equiv \boldsymbol{\sigma_R K_R^{-1} K_F \sigma_F}^{-1}, \tag{3.2.6}$$

where $\boldsymbol{\Phi}$ is an orthonormal basis transposer $\boldsymbol{K_R^{-1} K_F}$ that transposes a vector expressed in the space that spans the ex-ante forecast risk factor vectors $\boldsymbol{c_F}(1), \cdots,$ $\boldsymbol{c_F}(n)$, to a vector expressed in the space that spans the ex-post realized risk factor vectors $\boldsymbol{c_R}(1), \cdots, \boldsymbol{c_R}(n)$. Additionally, $\boldsymbol{\Phi}$ changes from a value per ex-ante forecast risks to a value per ex-post risks by $\boldsymbol{\sigma_F}^{-1}$ and $\boldsymbol{\sigma_R}$. Therefore, we call $\boldsymbol{K_R \sigma_R \Phi \sigma_F}^{-1}\boldsymbol{K_F^{-1} r_F}$ the risk basis transposed forecast return. Note that $\boldsymbol{K_F^{-1} r_F}$ is the ex-ante forecast factor return vector, and $\boldsymbol{\sigma_F}^{-1}\boldsymbol{K_F^{-1} r_F}$ is the ex-ante forecast factor Sharpe ratio vector. Thus, we have Theorem 3.1 (see Appendix 2.1).

Theorem 3.1:

The optimal portfolio $\boldsymbol{w}_F$, which is based on the ex-ante forecast return distribution $\boldsymbol{N}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F})$, is identical to the optimal portfolio $\boldsymbol{w}_0$ based on the return distribution $\boldsymbol{N}\left(\boldsymbol{K_R \sigma_R \Phi \sigma_F}^{-1}\boldsymbol{K_F^{-1} r_F}, \boldsymbol{\Sigma_R}\right)$ whose parameters are the risk basis transposed forecast return $\boldsymbol{K_R \sigma_R \Phi \sigma_F}^{-1}\boldsymbol{K_F^{-1} r_F}$ and the ex-post realized covariance matrix $\boldsymbol{\Sigma_R}$.

3.3 The Sharpe ratio achieved by an optimal portfolio

First, we consider the maximum Sharpe ratio of the optimal portfolio based on the ex-post realized return distribution $\boldsymbol{N}(\boldsymbol{r_R}, \boldsymbol{\Sigma_R})$. From (3.2.2), we have the optimal portfolio $\boldsymbol{w}_R$:

$$\boldsymbol{w}_R = \frac{\boldsymbol{\Sigma_R^{-1} r_R}}{\boldsymbol{e}'\boldsymbol{\Sigma_R^{-1} r_R}}. \tag{3.3.1}$$

The Sharpe ratio, which is achieved by the optimal portfolio previously described, is:

$$\mathrm{SR}_R = \frac{\boldsymbol{w}_R{}'\boldsymbol{r_R}}{\sqrt{\boldsymbol{w}_R{}'\boldsymbol{\Sigma_R w}_R}} = \frac{\boldsymbol{r_R}'\boldsymbol{\Sigma_R^{-1} r_R}}{\sqrt{\boldsymbol{r_R}'\boldsymbol{\Sigma_R^{-1} r_R}}} = \sqrt{\boldsymbol{r_R}'\boldsymbol{\Sigma_R^{-1} r_R}}. \tag{3.3.2}$$

From (3.2.4), we have $\boldsymbol{\Sigma_R^{-1}} = \boldsymbol{K_R \sigma_R^{-2} K_R^{-1}}$. Substituting it into (3.3.2), we have

$$\mathrm{SR}_R = \sqrt{\boldsymbol{r_R}'\boldsymbol{K_R \sigma_R^{-2} K_R^{-1} r_R}} = \sqrt{(\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R})'(\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R})} = \|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\|. \tag{3.3.3}$$

Note that $\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{r}_{\boldsymbol{R}}$ is the ex-post realized factor return vector, and $\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{r}_{\boldsymbol{R}}$ is the ex-post realized factor Sharpe ratio vector. Indeed, we find that $\left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{r}_{\boldsymbol{R}}\right\|$ indicates the magnitude of the absolute level of the realized market Sharpe ratio. Next, we consider the Sharpe ratio of an optimal portfolio based on a return distribution with parameters of an arbitrary return forecast $\boldsymbol{u}$ and ex-post realized covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{R}}$, where $\boldsymbol{u} \neq \boldsymbol{r}_{\boldsymbol{R}}$. As with the previous discussion, we have the optimal portfolio $\boldsymbol{w}_u$ based on the forecast distribution $\boldsymbol{N}(\boldsymbol{u}, \boldsymbol{\Sigma}_{\boldsymbol{R}})$

$$\boldsymbol{w}_u = \frac{\boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{u}}{\boldsymbol{e}'\boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{u}}. \qquad (3.3.4)$$

From (3.2.4), the Sharpe ratio from the optimal portfolio is

$$\mathrm{SR}_u = \frac{\boldsymbol{w}_u{}'\boldsymbol{r}_{\boldsymbol{R}}}{\sqrt{\boldsymbol{w}_u{}'\boldsymbol{\Sigma}_{\boldsymbol{R}}\boldsymbol{w}_u}} = \frac{\boldsymbol{u}'\boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{r}_{\boldsymbol{R}}}{\sqrt{\boldsymbol{u}'\boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{u}}} = \frac{\left(\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{u}\right)'\left(\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{r}_{\boldsymbol{R}}\right)}{\sqrt{\left(\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{u}\right)'\left(\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{u}\right)}}. \qquad (3.3.5)$$

Note that $\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{u}$ is the forecast factor return vector, and $\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{u}$ is the forecast factor Sharpe ratio vector, which is derived from $\boldsymbol{u}$. From (3.3.5), we have

$$\mathrm{SR}_u = \left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{r}_{\boldsymbol{R}}\right\|(\boldsymbol{h}, \boldsymbol{g}), \qquad (3.3.6)$$

where

$$\boldsymbol{h} \equiv \frac{\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{u}}{\left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{u}\right\|}, \boldsymbol{g} \equiv \frac{\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{r}_{\boldsymbol{R}}}{\left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{r}_{\boldsymbol{R}}\right\|}, \qquad (3.3.7)$$

and $(\boldsymbol{h}, \boldsymbol{g})$ is the inner product of $\boldsymbol{h}$ and $\boldsymbol{g}$. Note that $\boldsymbol{h}$ is the normalized forecast factor Sharpe ratio vector and that $\boldsymbol{g}$ is the normalized ex-post realized factor Sharpe ratio vector.

Suppose we have two return forecasts, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. Let SR_{u_1} and SR_{u_2} represent the Sharpe ratios from the return forecasts $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. We then calculate $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ from (3.3.7). From (3.3.6), we obtain

$$\mathrm{SR}_{u_1} \leq \mathrm{SR}_{u_2} \Leftrightarrow (\boldsymbol{h}_1, \boldsymbol{g}) \leq (\boldsymbol{h}_2, \boldsymbol{g}). \qquad (3.3.8)$$

Assuming that an investor achieves a higher Sharpe ratio when he or she has a stronger forecast is natural. Thus, we derive Property 3.1.

Property 3.1:

The inner product $(\boldsymbol{h}, \boldsymbol{g})$ is a measure for evaluating the power of the return forecast $\boldsymbol{u}$.

We calculate the factor Sharpe ratio vector derived from the risk basis transposed forecast return $\boldsymbol{K}_{\boldsymbol{R}}\boldsymbol{\sigma}_{\boldsymbol{R}}\boldsymbol{\Phi}\boldsymbol{\sigma}_{\boldsymbol{F}}{}^{-1}\boldsymbol{K}_{\boldsymbol{F}}^{-1}\boldsymbol{r}_{\boldsymbol{F}}$ as follows:

$$\boldsymbol{\sigma}_{\boldsymbol{R}}{}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\left(\boldsymbol{K}_{\boldsymbol{R}}\boldsymbol{\sigma}_{\boldsymbol{R}}\boldsymbol{\Phi}\boldsymbol{\sigma}_{\boldsymbol{F}}{}^{-1}\boldsymbol{K}_{\boldsymbol{F}}^{-1}\boldsymbol{r}_{\boldsymbol{F}}\right) = \boldsymbol{\Phi}\boldsymbol{\sigma}_{\boldsymbol{F}}{}^{-1}\boldsymbol{K}_{\boldsymbol{F}}^{-1}\boldsymbol{r}_{\boldsymbol{F}}. \qquad (3.3.9)$$

Therefore, the normalized factor Sharpe ratio vector $\boldsymbol{f}$ derived from the risk basis transposed forecast return $\boldsymbol{K_R \sigma_R \Phi \sigma_F}^{-1}\boldsymbol{K_F^{-1} r_F}$ is:

$$\boldsymbol{f} = \frac{\boldsymbol{\Phi}\left(\boldsymbol{\sigma_F}^{-1}\boldsymbol{K_F^{-1} r_F}\right)}{\left\|\boldsymbol{\Phi}\left(\boldsymbol{\sigma_F}^{-1}\boldsymbol{K_F^{-1} r_F}\right)\right\|}. \tag{3.3.10}$$

From (3.3.6), we derive Property 3.2.

Property 3.2:

The Sharpe ratio $\mathrm{SR_F}$, which is the optimal portfolio based on the ex-ante forecast return distribution $\boldsymbol{N(r_F, \Sigma_F)}$, is expressed by the following:

$$\mathrm{SR_F} = \left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\|(\boldsymbol{f}, \boldsymbol{g}). \tag{3.3.11}$$

Given that $\boldsymbol{f}$ includes information on the ex-ante forecast return and the difference between the ex-ante forecast covariance matrix and the ex-post realized covariance matrix, we derive Property 3.3.

Property 3.3:

The inner product of the normalized factor Sharpe ratio derived from the risk basis transposed ex-ante forecast return and the normalized ex-post realized factor Sharpe ratio $(\boldsymbol{f}, \boldsymbol{g})$ is a measure that simultaneously evaluates the ex-ante forecast risk and return. Subsequently, we refer to the measure as *the predictability-based measurement*.

From (3.3.11), we show that the realized Sharpe ratio is the product of the introduced measure multiplied by the absolute level of the realized market Sharpe ratio. Because $\boldsymbol{f}$ and $\boldsymbol{g}$ are both normalized, the measure is a number between one and negative one, which does not include information on the absolute value of the market Sharpe ratio. We can separately evaluate the ex-ante forecast from the market conditions using the measure.

3.4 Numerical examples of the measurement

We calculate the predictability-based measurement using a historical data sample. We use an investment of ten weekly S&P sector total return indices.

First, we need the sample forecast return distributions that are evaluated by the predictability-based measurement. We use the forecast distributions that have some meaning instead of arbitrary distributions. Note that the forecast return distribution has the forecast return vector and the forecast covariance matrix. First, we assume that a test forecast covariance matrix for the next year at the end of the current year is identical to the realized covariance matrix of the current year. Second, remember that we can construct an optimal portfolio from the given forecast return vector and the given risk model using the mean-variance approach under modern portfolio theory. Then, for a given portfolio, forecast return vectors lead the portfolio as the optimal portfolio using the mean-variance approach and the test forecast risk. In this paper, we call these forecast return vectors the implied forecast returns. We can then calculate the implied forecast return of a given portfolio as the return vector, which gives us the mean-variance optimal portfolio (see Appendix 2.2).

We use the market weight portfolio, the equal weight portfolio, and the minimum variance portfolio as the three given portfolios as of the end of each year from 2010 through 2014 (see derivations of the minimum variance portfolio weight in Appendix 2.3). Second, we calculate the implied forecast return of the three portfolios at the end of each year using the mean-variance approach and the test forecast risk. These are derived from the three portfolio weights and the test forecast risk, which is in fact the realized risk of the current year. Thus, we have three sets of risk and return forecasts at the end of each year. Additionally, we note that the equal-weight portfolio has an interesting property that its performance is identical to the theoretical average of all investment strategies (see Appendix 3).

Let $\boldsymbol{r}_{F\_Mkt}(t)$, $\boldsymbol{r}_{F\_Eq}(t)$, and $\boldsymbol{r}_{F\_Min}(t)$ denote the forecast return vector for year t implied by the market weight portfolio, the equal weight portfolio, and the minimum variance portfolio, respectively, at the end of year $t-1$. Furthermore, let $\boldsymbol{\Sigma}_F(t)$ denote the forecast risk for year t at the end of year $t-1$, and let $\boldsymbol{\Sigma}_R(t)$ denote the realized risk during year t. Indeed, $\boldsymbol{\Sigma}_R(t)$ is the covariance calculated from the fifty-two weekly returns of the ten weekly S&P sector total return indices during year t. Using the test forecast risk, $\boldsymbol{\Sigma}_F(t) \equiv \boldsymbol{\Sigma}_R(t-1)$.

For example, we calculate the forecast distribution implied by the market weight portfolio F_Mkt, the equal weight portfolio F_Eq, and the minimum variance portfolio F_Min at the end of 2010 as follows:

$$\text{F\_Mkt:} \quad \boldsymbol{N}\left(\boldsymbol{r}_{F\_Mkt}(2011), \boldsymbol{\Sigma}_{\boldsymbol{F}}(2011)\right), \tag{3.4.1}$$

$$\text{F\_Eq:} \quad \boldsymbol{N}\left(\boldsymbol{r}_{F\_Eq}(2011), \boldsymbol{\Sigma}_{\boldsymbol{F}}(2011)\right), \tag{3.4.2}$$

$$\text{F\_Min:} \quad \boldsymbol{N}\left(\boldsymbol{r}_{F\_Min}(2011), \boldsymbol{\Sigma}_{\boldsymbol{F}}(2011)\right). \tag{3.4.3}$$

Let $\boldsymbol{w}_{Mkt}(t), \boldsymbol{w}_{Eq}(t),$ and $\boldsymbol{w}_{Min}(t)$ denote the weight vectors of the market portfolio, equal weight portfolio, and minimum variance portfolio, respectively, at the end of year t. From the test forecast risk, $\boldsymbol{\Sigma_F}(2011) = \boldsymbol{\Sigma}_R(2010)$. Additionally, we have $\boldsymbol{r}_{F\_Mkt}(2011) = \boldsymbol{\Sigma}_F(2011)\boldsymbol{w}_{Mkt}(2010) = \boldsymbol{\Sigma}_R(2010)\boldsymbol{w}_{Mkt}(2010)$ (Appendix 3).

Table 3.1 indicates the Sharpe ratios for one-year buy-and-hold investments of the three portfolios without friction such as transaction costs for the years 2011, 2012, 2013, 2014, and 2015. These ratios also indicate the investment results of the optimal portfolios from the three implied forecasts that maximize the Sharpe ratio using the mean-variance approach.

Table 3.1 Predictability-based measurement of the three implied forecasts

Year t	Sharpe ratio of the portfolio			Market condition $\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{r}_{\boldsymbol{R}}\|$	Predictability-based measurement of implied forecast at the end of the previous year $(\boldsymbol{f},\boldsymbol{g})$		
	Market weight	Equal weight	Minimum variance		F_Mkt(1)	F_Eq(2)	F_Min(3)
2015	-0.000	-0.022	0.030	0.375	-0.001	-0.058	0.081
2014	0.195	0.194	0.309	0.671	0.290	0.289	0.461
2013	0.475	0.438	0.348	0.853	0.557	0.514	0.408
2012	0.168	0.168	-0.014	0.403	0.415	0.416	-0.035
2011	0.014	0.026	0.187	0.504	0.028	0.053	0.371

(1) F_Mkt: Implied forecast based on the market weight portfolio at the end of the previous year.
(2) F_Eq: Implied forecast based on the equal weight portfolio at the end of the previous year.
(3) F_Min: Implied forecast based on the minimum variance portfolio at the end of the previous year.

The implied forecast from the market weight portfolio at the end of 2012 achieved the highest Sharpe ratio for the next year, or 0.475. The implied forecast from the equal weight portfolio at the end of 2014 achieved the lowest Sharpe ratio of the next year, or –0.022. Additionally, the best forecast is the implied forecast from the market weight portfolio at the end of 2012. The predictability-based measurement of the implied forecast is 0.557. The worst forecast is the implied forecast from the equal weight portfolio at the end of 2014. The predictability-based measurement is –0.058. The best of these forecasts achieves the highest Sharpe ratio, and the worst achieves the lowest Sharpe ratio, which is natural.

However, we find some interesting cases. We compare the predictability-based measurement of the implied forecasts based on the market weight portfolio at the end of 2011 with the measure of the implied forecasts based on the market weight portfolio at the end of 2013. The former is 0.415, and the latter is 0.290. However, the Sharpe ratio of the market weight portfolio at the end of 2013 is 0.195 for the next year, whereas the Sharpe ratio of the market weight portfolio at the end of 2011 is 0.168 for the next year. Thus, the Sharpe ratio in 2014 is relatively higher than that in 2012 despite the lower value of the predictability-based measurement. This result comes from the difference in market conditions between the two years. We recall that the market conditions indicate the magnitude of the absolute level of the Sharpe ratio $\left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-1}\boldsymbol{r}_{\boldsymbol{R}}\right\|$, a number that reflects the economic environment. For example, a lower number in 2015 is recognized as a reflection of the China shock. The value in 2014 is 0.671, and the value in 2012 is 0.403. Therefore, the market conditions affect the previously described situation. Other similarities exist between these forecasts, such as the pair of implied forecasts based on the equal weight portfolios at the end of 2011 and 2013. The predictability-based measurement of the implied forecasts based on the former is 0.416 and based on the latter is 0.289. However, the Sharpe ratio of the equal weight portfolio at the end of 2013 is 0.194 for the next year, whereas the Sharpe ratio of the equal weight portfolio at the end of 2011 is 0.168 for the next year. We find one more case, which is the pair of implied forecasts based on the minimum variance portfolio at the end of 2012 and 2013. The predictability-based measurement of the implied forecasts based on the former is 0.408 and based on the latter is 0.461. However, the Sharpe ratio of the minimum variance portfolio at the end of 2012 is 0.348 for the next year and at the end of 2013 is 0.309 for the next year. From these cases, we recognize the importance of evaluating the forecast separately from the market conditions.

3.5 Sharpe Ratio for multi-period investments

In investment practices, investors often rebalance their portfolio by every fixed period, for example monthly, that each period is adequate for the investment. Indeed, they re-organize their portfolio based on their same forecasts unless their forecasts are changed. Additionally, investment results during multi periods are usually analyzed for performance evaluation for the investment. In this section, discussions are extended for multi periods investment.

Suppose a simple case of multi periods investment that there are T equal periods and ex-post realized returns are not change unless the ex-ante forecast returns are change. Suppose that the return $\boldsymbol{r}$ of the n risky assets in an investment period has the normal distribution $\boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and there are continuing T investment periods of this. Indeed, suppose that the risky assets have return $\boldsymbol{r}(t)$ in the $t-th$ period and that the returns $\boldsymbol{r}(t)$ are independent of each other and identically distributed (we use the notation "idd").

The observed Sharpe ratio during whole T periods is calculated using the sample returns of $\boldsymbol{r}(t), t = 1, \cdots, T$. For example, we calculate the Sharpe ratio using the monthly returns of a given investment result to evaluate the quality of the investment. In this case, we have $T = 12$.

Suppose we have a portfolio $\boldsymbol{w}$ at the beginning of each investment period because of the assumption that the ex-ante forecasts are identical in each investment period. After each period, the investment results are

$$\boldsymbol{w}'\boldsymbol{r}(t), \tag{3.5.1}$$

because of the assumption that the ex-post realized returns are not change unless the ex-ante forecast returns are change. Therefore, we have the average return during whole T periods

$$\frac{1}{T}\sum_t \boldsymbol{w}'\boldsymbol{r}(t), \tag{3.5.2}$$

and the unbiased estimator of standard deviation from sample returns

$$\sqrt{\frac{\sum_t \left(\boldsymbol{w}'\boldsymbol{r}(t) - \frac{1}{T}\sum_t \boldsymbol{w}'\boldsymbol{r}(t)\right)^2}{T-1}}. \tag{3.5.3}$$

From $\boldsymbol{r}(t) \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$,

$$\frac{\sum_t \left(\boldsymbol{w}'\boldsymbol{r}(t) - \frac{1}{T}\sum_t \boldsymbol{w}'\boldsymbol{r}(t)\right)^2}{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}} \sim \boldsymbol{\chi}^2(T-1), \tag{3.5.4}$$

where $\boldsymbol{\chi}^2(m)$ is the chi-squared distribution with m degrees of freedom.
Therefore, the observed Sharpe ratio during the whole period $\mathrm{SR_{ob}}$:

$$
\begin{aligned}
\mathrm{SR_{ob}} &\equiv \frac{\frac{1}{T}\sum_t \boldsymbol{w}'\boldsymbol{r}(t)}{\sqrt{\frac{\sum_t\left(\boldsymbol{w}'\boldsymbol{r}(t)-\frac{1}{T}\sum_t \boldsymbol{w}'\boldsymbol{r}(t)\right)^2}{T-1}}} \\
&= \frac{\frac{1}{T}\sum_t \boldsymbol{w}'\boldsymbol{r}_T(t)}{\sqrt{\frac{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}}{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}}\frac{\sum_t\left(\boldsymbol{w}'\boldsymbol{r}_T(t)-\frac{1}{\mathrm{T}}\sum_t \boldsymbol{w}'\boldsymbol{r}_T(t)\right)^2}{T-1}}} \\
&= \frac{\frac{\frac{1}{T}\sum_t \boldsymbol{w}'\boldsymbol{r}_T(t)}{\sqrt{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}}}}{\sqrt{\frac{\frac{1}{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}}\sum_t\left(\boldsymbol{w}'\boldsymbol{r}_T(t)-\frac{1}{T}\sum_t \boldsymbol{w}'\boldsymbol{r}_T(t)\right)^2}{T-1}}}
\end{aligned}
\tag{3.5.5}
$$

has the noncentral $\boldsymbol{t}$ distribution with $T-1$ degrees of freedom and noncentrality parameter $\delta = \boldsymbol{w}'\boldsymbol{\mu}/\sqrt{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}}$.

Finally, the expectation value of the observable Sharpe ratio is the first-order moment of the distribution noncentral $\boldsymbol{t}$ distribution with $T-1$ degrees of freedom and noncentrality parameter $\delta = \boldsymbol{w}'\boldsymbol{\mu}/\sqrt{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}}$:

$$
\mathrm{E}[\mathrm{SR_{ob}}] = \frac{\delta\left(\frac{T}{2}\right)^{1/2}\boldsymbol{\Gamma}\left(\frac{T-1}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{T}{2}\right)}, \tag{3.5.6}
$$

where $\boldsymbol{\Gamma}(n)$ is the Gamma function. Note that the noncentrality parameter is the ex-post Sharpe ratio achieved by the optimal portfolio based on the ex-ante forecast return.

Typically, we use monthly return data in the performance analysis. Additionally, we sometimes have time series of returns that have different time horizon lengths. Therefore, we observe the difference in the observable Sharpe ratio measured by the sample monthly returns relative to the Sharpe ratio of the return distribution. Again, note that the returns $\boldsymbol{r}_T(t)$ are iid.

Suppose that we have a Sharpe ratio measured by some length of investment periods. The noncentrality parameter is the ex-post Sharpe ratio. Therefore, the expectation value of the observed Sharpe ratio is the product of the ex-post Sharpe ratio and the coefficient from (3.5.6). Numerical calculations of the coefficient:

$$
\frac{\left(\frac{T}{2}\right)^{1/2}\boldsymbol{\Gamma}\left(\frac{T-1}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{T}{2}\right)}, \tag{3.5.7}
$$

for some cases are following.

Case of $T = 6$, half a year,

$$\frac{\left(\frac{6}{2}\right)^{1/2}\boldsymbol{\Gamma}\left(\frac{6-1}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{6}{2}\right)} = \frac{(3)^{1/2}\boldsymbol{\Gamma}\left(\frac{5}{2}\right)}{\boldsymbol{\Gamma}(3)} = \frac{\sqrt{3}\frac{3}{2}\frac{1}{2}\boldsymbol{\Gamma}\left(\frac{1}{2}\right)}{2\cdot 1\cdot\boldsymbol{\Gamma}(1)} = \sqrt{3}\frac{\frac{3}{2}\frac{1}{2}}{2\cdot 1}\sqrt{\pi}$$

$$= \sqrt{3}\sqrt{\pi}\frac{3}{2\cdot 2\cdot 2} \cong 1.151. \tag{3.5.8}$$

Cases of $T = 6, 8,$ and 10,

$$\frac{\left(\frac{6}{2}\right)^{1/2}\boldsymbol{\Gamma}\left(\frac{6-1}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{6}{2}\right)} = \frac{(3)^{1/2}\boldsymbol{\Gamma}\left(\frac{5}{2}\right)}{\boldsymbol{\Gamma}(3)} = \frac{\sqrt{3}\frac{3}{2}\frac{1}{2}\boldsymbol{\Gamma}\left(\frac{1}{2}\right)}{2\cdot 1\cdot\boldsymbol{\Gamma}(1)} = \sqrt{3}\frac{\frac{3}{2}\frac{1}{2}}{2\cdot 1}\sqrt{\pi}$$

$$= \sqrt{3}\sqrt{\pi}\frac{3}{2\cdot 2\cdot 2} \cong 1.151, \tag{3.5.9}$$

$$\frac{\left(\frac{8}{2}\right)^{1/2}\boldsymbol{\Gamma}\left(\frac{8-1}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{8}{2}\right)} = \frac{(4)^{1/2}\boldsymbol{\Gamma}\left(\frac{7}{2}\right)}{\boldsymbol{\Gamma}(4)} = \frac{\sqrt{4}\frac{5}{2}\frac{3}{2}\frac{1}{2}\boldsymbol{\Gamma}\left(\frac{1}{2}\right)}{3\cdot 2\cdot 1\cdot\boldsymbol{\Gamma}(1)} = \sqrt{4}\frac{\frac{5}{2}\frac{3}{2}\frac{1}{2}}{3\cdot 2\cdot 1}\sqrt{\pi}$$

$$= \sqrt{4}\sqrt{\pi}\frac{5}{2\cdot 2\cdot 2\cdot 2} \cong 1.108, \text{and} \tag{3.5.10}$$

$$\frac{\left(\frac{10}{2}\right)^{1/2}\boldsymbol{\Gamma}\left(\frac{10-1}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{10}{2}\right)} = \frac{(5)^{1/2}\boldsymbol{\Gamma}\left(\frac{9}{2}\right)}{\boldsymbol{\Gamma}(5)} = \frac{\sqrt{5}\frac{7}{2}\frac{5}{2}\frac{3}{2}\frac{1}{2}\boldsymbol{\Gamma}\left(\frac{1}{2}\right)}{4\cdot 3\cdot 2\cdot 1\cdot\boldsymbol{\Gamma}(1)} = \sqrt{5}\frac{\frac{7}{2}\frac{5}{2}\frac{3}{2}\frac{1}{2}}{4\cdot 3\cdot 2\cdot 1}\sqrt{\pi}$$

$$= \sqrt{5}\sqrt{\pi}\frac{7\cdot 5}{4\cdot 2\cdot 2\cdot 2\cdot 2\cdot 2} \cong 1.084. \tag{3.5.11}$$

We summarize the sample results as follows (more samples in Table A4.1 in Appendix 4).

Table 3.2: Overestimation of observed Sharpe ratio

The number of periods	The coefficient
12	1.0684
24	1.0327
36	1.0215
60	1.0127

Again, we suppose that there are some lengths of equal periods and ex-post realized returns are not change unless the ex-ante forecast returns are change. Indeed, this means that the ex-post realized returns $\boldsymbol{r}(t)$ are iid during the periods. Under the

assumption, the observed Sharpe ratio converges to the ex-post realized Sharpe ratio achieved by optimal portfolio based on the ex-ante forecast when the number of periods becomes longer (see Appendix 4). The observed Sharpe ratio shows an over value of the ex-post realized Sharpe ratio. Additionally, the over value becomes smaller when the length longer. Therefore, the observed Sharpe ratio measured by a small number of periods may overestimate the potential Sharpe ratio that the portfolio achieves.
For example, if the samples are monthly returns, we may have an almost seven percent overestimation of the Sharpe ratio for twelve months of data, three percent overestimation for two years of data, and two percent overestimation for three years of data.

3.6 Summary

In this chapter, we offer the predictability-based measurement that evaluates the ex-ante forecast risk and return. We also show that the predictability-based measurement evaluates an ex-ante forecast separate from market conditions. Finally, we show that the Sharpe ratio that reflects the optimal portfolio achieved by the forecast is the product of the predictability-based measurement multiplied by the absolute value of the market Sharpe ratio.

We find that the normalized ex-ante forecast Sharpe ratios are crucial to the investment results. This finding indicates that conducting additional relative analysis to discern the attractiveness of the investable assets is important. In addition, we recognize the importance of both the risk and the return forecasts. We observe many practical investment cases in which return forecasts are the top priority, followed by risk forecasts. However, our discussion indicates that for investors to develop and use in-house risk models that forecast the risk factors and their risks with a high degree of accuracy may be advantageous.

In addition, the predictability-based measurement that evaluates forecasts separate from the market conditions may become a useful analytical tool for portfolio managers and sponsors.

Additional research should be performed on the applications of the properties discussed in this paper. For example, Property 3.2 shows that multiple strong forecasts achieve a high Sharpe ratio, which may indicate the possibility of multiple methods that create strong forecasts. Additionally, we observe how changes in risk factors and their risks affect forecast quality. Property 3.2 provides a better understanding of the Sharpe ratio, which may facilitate a more practical analysis. Finally, further research is required to determine the effectiveness of the properties from different aspects of the investment process.

Chapter 4

Attribution of ex-post realized Sharpe ratio to predictability-based measurement of ex-ante forecast returns and risk

4.1 Introduction

As was mentioned in the previous chapter, improving the Sharpe ratio is an important goal of portfolio managers and sponsors. This Sharpe ratio is based on portfolio performance under an ex-post realized return distribution. However, investors construct portfolios based on ex-ante forecast return distributions for investable assets at the beginning of an investment period by using the mean-variance approach. As such, it is natural that the ex-ante forecast return distribution differs from the ex-post return distributions and realized future distributions. Investors recognize that attribution analysis of past investments is important for their future investments. Therefore, we need to attribute the Sharpe ratio to the ex-ante forecast return distribution after the investment period.

In this chapter, brief reviews are again provided for the predictability-based measurement for the Sharpe ratio as introduced in the previous chapter. Additionally, the realized Sharpe ratio is decomposed into the market condition, the predictability-based measurement of returns, and the predictability-based measurement of risk, which is further decomposed into the predictability-based measurement of magnitude of risk and the predictability-based measurement of risk factors.

Regarding Sharpe ratio attribution analysis, Steiner (2011) decomposes the Sharpe ratio of a portfolio into risk weights, diversification effects, and asset Sharpe ratios, as reviewed in Chapter 2. The risk weight is defined as $w_i \rho_{i,p} \sigma_i / \sigma_p$, where w_i is the weight of the ith asset in portfolio p, $\rho_{i,p}$ the asset covariance with the overall portfolio, σ_i the volatility of the asset, and σ_p the volatility of the portfolio. However, ex-ante forecast returns and risk lead to the weight of ith asset of the portfolio in the mean-variance approach. It is thus important to analyze the sources of the input that lead the weight of ith asset of the portfolio. As such, a Sharpe ratio decomposition that directly uses the forecast returns and risk instead of the weights is discussed.

The first contribution of this chapter is proposing an ex-post realized Sharpe ratio attribution to ex-ante forecast returns and risk. The second contribution is providing the formula for the attribution. To this end, we decompose the predictability-based measurement of risk into the predictability-based measurement of magnitude of risk and the predictability-based measurement of risk factors. This attribution formula enables us to compare predictability-based measurement of return and predictability-based measurement of risk, in a quantitative way. The third contribution is identifying the excess Sharpe ratio attribution against a given benchmark using the benchmark portfolio's predictability-based measurement of the implied returns forecast. As a result, we find the relative strength of the optimal portfolio against the benchmark, as well as the absolute strengths of the portfolio and benchmark. Consequently, our attribution analysis can be applied to diverse situations and to derive rich information.

We here illustrate a simple example of a Sharpe ratio attribution analysis which is similar as in the previous chapter. We assume that we have three investable assets: Asset A, Asset B, and Asset C. We assume that Asset A has ex-post returns of $r_A = 6\%$, Asset B of $r_B = 4\%$, and Asset C of $r_C = 1\%$, respectively. Regarding their ex-post risks of returns, the three assets have ex-post standard deviations $\sigma_A = 8\%, \sigma_B = 8\%,$ and $\sigma_C = 2\%$, the returns of Asset A and Asset B have the correlation $\rho_{AB} = 0.6$, the returns of Asset A and Asset C are independent, and the returns of Asset B and Asset C are also independent. Given this return distribution, we can construct the optimal portfolio, which is 25% long Asset A, 3% long Asset B, and 72% long Asset C. The optimal portfolio has a maximum Sharpe ratio of 0.90. Suppose that an investor has a perfect ex-ante return forecast and an imperfect ex-ante risk forecast. We assume that the investor forecasts that Asset A has an ex-ante return of $r_A = 6\%$, Asset B of $r_B = 4\%$, and Asset C of $r_C = 1\%$, respectively. Additionally, we assume that the investor forecasts the ex-ante standard deviations of $\sigma_A = 10\%, \sigma_B = 6\%,$ and $\sigma_C = 2\%$ while pairs of returns are independent. However, the investor must construct the optimal portfolio from his or her ex-ante forecast return distribution using the mean-variance approach as 14% long on Asset A, 26% long on Asset B, and 59% long on Asset C. The optimal portfolio achieves a Sharpe ratio of 0.79 under the ex-post realized return distribution. The optimal portfolio delivers a Sharpe ratio that is 87% of the maximum Sharpe ratio of the optimal portfolio from the ex-post realized distribution. We first decompose the realized Sharpe ratio of the optimal portfolio of 0.79 into the product of the realized market conditions of 0.90, which is the achievable maximum Sharpe ratio by an investor, and into the predictability-based measurement of ex-ante forecast of 0.87, based on Shimizu (2017). Additionally, we decompose the predictability-based measurement of ex-ante forecast into the predictability-based measurement of ex-ante return forecast of 1.0 and into a predictability-based measurement of ex-ante risk forecast of 0.87. It is natural that the predictability-based measurement of ex-ante return forecast is 1.0 because the investor has perfect return forecasts. Furthermore, we decompose the predictability-based measurement of risk of 0.87 into a predictability-based measurement of the magnitude of risk of almost 1.0 and a predictability-based measurement of the risk factors of 0.82. From the above example, the realized Sharpe ratio reduction is caused by weak risk predictability despite the perfect return predictability.

This chapter is organized as follows. In Section 4.2, we determine that the ex-post realized Sharpe ratio is attributed to the realized market conditions, predictability-based measurement of returns, predictability-based measurement of risk magnitude, and predictability-based measurement of risk factors. In Section 4.3, we provide numerical examples of the attribution analysis using an investment of ten weekly S&P sector total return indices.

In Section 4.4, we discuss the excess Sharpe ratio attribution against a given benchmark. Assuming the implied return forecast of the benchmark portfolio, we decompose the excess Sharpe ratio into the effect of the predictability-based measurements of return and that of the risk. Additionally, we provide numerical examples for the excess Sharpe ratio attribution analysis, assuming the investor's optimal portfolio is the equal weight portfolio and the benchmark portfolio is the market weight one. We observe rich information from the attribution analysis.

*An outline and results of this chapter of the thesis is described below because of including the contents published in the future in Shimizu, M. (Forthcoming).

4.2 Attribution of Sharpe ratio

We consider a one-period investment in a non-leveraged long/short portfolio. We have n risky assets and construct the portfolio using the mean-variance approach without leverage to maximize the expected Sharpe ratio. We suppose that the returns of the n risky assets behave according to an n-dimensional normal distribution $\boldsymbol{N}(\boldsymbol{r},\boldsymbol{\Sigma})$, where $\boldsymbol{r}$ is a vector of the returns of the risky assets and $\boldsymbol{\Sigma}$ is a covariance matrix of the returns. We also suppose that the covariance matrix of the n-dimensional normal distribution has a full rank.

Denote $SR(\boldsymbol{r_F},\boldsymbol{\Sigma_F} \mid \boldsymbol{r_R},\boldsymbol{\Sigma_R})$ as the ex-post realized Sharpe ratio that the optimal portfolio from the ex-ante forecast return distribution $\boldsymbol{N}(\boldsymbol{r_F},\boldsymbol{\Sigma_F})$ achieves under the ex-post realized return distribution $\boldsymbol{N}(\boldsymbol{r_R},\boldsymbol{\Sigma_R})$. From Property 3.2, we have

$$SR(\boldsymbol{r_F},\boldsymbol{\Sigma_F} \mid \boldsymbol{r_R},\boldsymbol{\Sigma_R}) = \left\|\boldsymbol{\sigma_R}^{-1}\boldsymbol{K_R}^{-1}\boldsymbol{r_R}\right\|\left(\frac{\boldsymbol{\Phi}\boldsymbol{\sigma_F}^{-1}\boldsymbol{K_F}^{-1}\boldsymbol{r_F}}{\left\|\boldsymbol{\Phi}\boldsymbol{\sigma_F}^{-1}\boldsymbol{K_F}^{-1}\boldsymbol{r_F}\right\|},\boldsymbol{g}\right), \tag{4.2.1}$$

where

$$\boldsymbol{\Phi} \equiv \boldsymbol{\sigma_R}\boldsymbol{K_R}^{-1}\boldsymbol{K_F}\boldsymbol{\sigma_F}^{-1}, \boldsymbol{g} \equiv \frac{\boldsymbol{\sigma_R}^{-1}\boldsymbol{K_R}^{-1}\boldsymbol{r_R}}{\left\|\boldsymbol{\sigma_R}^{-1}\boldsymbol{K_R}^{-1}\boldsymbol{r_R}\right\|}, \tag{4.2.2}$$

and $(\boldsymbol{x},\boldsymbol{y})$ is the inner product of vector $\boldsymbol{x}$ and vector $\boldsymbol{y}$ and $\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x},\boldsymbol{x})}$ the norm of $\boldsymbol{x}$.

Recall that the inner product is the predictability-based measurement of ex-ante forecast returns and risk. We introduce the quantity $P\_SR(\boldsymbol{r_F},\boldsymbol{\sigma_F},\boldsymbol{K_F})$, which has three parameters, $\boldsymbol{r_F},\boldsymbol{\sigma_F}$, and $\boldsymbol{K_F}$, for the ex-ante forecast distribution defined by the inner product between the normalized vector of $\boldsymbol{\Phi}\boldsymbol{\sigma_F}^{-1}\boldsymbol{K_F}^{-1}\boldsymbol{r_F}$ and $\boldsymbol{g}$, that is,

$$P\_SR(\boldsymbol{r_F},\boldsymbol{\sigma_F},\boldsymbol{K_F}) \equiv \left(\frac{\boldsymbol{\Phi}\boldsymbol{\sigma_F}^{-1}\boldsymbol{K_F}^{-1}\boldsymbol{r_F}}{\left\|\boldsymbol{\Phi}\boldsymbol{\sigma_F}^{-1}\boldsymbol{K_F}^{-1}\boldsymbol{r_F}\right\|},\boldsymbol{g}\right). \tag{4.2.3}$$

Indeed, $\boldsymbol{P}(\boldsymbol{r_F},\boldsymbol{\sigma_F},\boldsymbol{K_F})$ measures the predictability-based measurement of the overall ex-ante forecast return and risk. Substituting $\boldsymbol{\sigma_F}$ and $\boldsymbol{K_F}$ with $\boldsymbol{\sigma_R}$ and $\boldsymbol{K_R}$, respectively, into (4.2.3), we get:

$$P\_SR(\boldsymbol{r_F}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) \equiv \left(\frac{\boldsymbol{\Phi}\boldsymbol{\sigma_R}^{-1}\boldsymbol{K_R^{-1}}\boldsymbol{r_F}}{\left\|\boldsymbol{\Phi}\boldsymbol{\sigma_R}^{-1}\boldsymbol{K_R^{-1}}\boldsymbol{r_F}\right\|}, \boldsymbol{g}\right) = \left(\frac{\boldsymbol{\sigma_R}^{-1}\boldsymbol{K_R^{-1}}\boldsymbol{r_F}}{\left\|\boldsymbol{\sigma_R}^{-1}\boldsymbol{K_R^{-1}}\boldsymbol{r_F}\right\|}, \boldsymbol{g}\right). \qquad (4.2.4)$$

Similarly, we define the predictability-based measurement for the Sharpe ratio of the ex-ante forecast risk magnitude that is not identical to the ex-post realized risk magnitude

$$P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_R}) \equiv \left(\frac{\boldsymbol{\sigma_R}\boldsymbol{\sigma_F^{-2}}\boldsymbol{K_R^{-1}}\boldsymbol{r_R}}{\left\|\boldsymbol{\sigma_R}\boldsymbol{\sigma_F^{-2}}\boldsymbol{K_R^{-1}}\boldsymbol{r_R}\right\|}, \boldsymbol{g}\right), \qquad (4.2.5)$$

and the predictability-based measurement for the Sharpe ratio of the ex-ante forecast risk factors that is not identical to the ex-post realized risk factors

$$P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_F}) \equiv \left(\frac{\boldsymbol{\sigma_R}\boldsymbol{K_R^{-1}}\boldsymbol{K_F}\boldsymbol{\sigma_R}^{-2}\boldsymbol{K_F^{-1}}\boldsymbol{r_R}}{\left\|\boldsymbol{\sigma_R}\boldsymbol{K_R^{-1}}\boldsymbol{K_F}\boldsymbol{\sigma_R}^{-2}\boldsymbol{K_F^{-1}}\boldsymbol{r_R}\right\|}, \boldsymbol{g}\right). \qquad (4.2.6)$$

Again, the vectors in these three inner products are normalized; therefore, their values are numbers between one and negative one. Then, we find that the value

$$1 - P\_SR(\boldsymbol{r_F}, \boldsymbol{\sigma_F}, \boldsymbol{K_F}) \qquad (4.2.7)$$

shows the reduction in the Sharpe ratio that occurs because the ex-ante forecast distribution is not identical to the ex-post realized distribution.

We define a quantity DTY_SR expressing the synergy effect:

$$DTY\_SR$$
$$\equiv \left(1 - \boldsymbol{P}(\boldsymbol{r_F}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})\right) - \left\{\left(1 - \boldsymbol{P}(\boldsymbol{r_F}, \boldsymbol{\sigma_R}, \boldsymbol{K_R})\right) + \left(1 - \boldsymbol{P}(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})\right)\right\}. \qquad (4.2.8)$$

We call DTY_SR duplicate term Y. Therefore, we have

$$1 - P\_SR(\boldsymbol{r_F}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})$$
$$= \left(1 - P\_SR(\boldsymbol{r_F}, \boldsymbol{\sigma_R}, \boldsymbol{K_R})\right) + \left(1 - P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})\right) + DTY\_SR. \qquad (4.2.9)$$

Similarly, we have

$$1 - P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})$$
$$= \left(1 - P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_R})\right) + \left(1 - P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_F})\right) + DTX\_SR, \qquad (4.2.10)$$

where DTX_SR is the duplicate term X for the Sharpe ratio. We finally derive Property 4.1.

Property 4.1:

The ex-post realized Sharpe ratio is attributed to six components. One is the absolute level of the ex-post realized Sharpe ratio, $\left\|\boldsymbol{\sigma_R^{-1}}\boldsymbol{K_R^{-1}}\boldsymbol{r_R}\right\|$, which represents the market conditions over the investment period. The next three components are the predictability-based measurement for the Sharpe ratio of the ex-ante forecast return, $\boldsymbol{P}(\boldsymbol{r_F}, \boldsymbol{\sigma_R}, \boldsymbol{K_R})$, the ex-ante forecast magnitude of risk, $\boldsymbol{P}(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_R})$, and the ex-ante forecast risk factors, $\boldsymbol{P}(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_F})$, which an investor uses to construct an optimal portfolio with the mean-variance approach at the beginning of the period. The others are two duplicate terms,

DTX_SR and DTY_SR, for the Sharpe ratio.

4.3 Numerical examples of attributions

In this section, we demonstrate our attribution analysis using a historical data sample. We use an investment of weekly ten S&P500 sector total return indices for 471 weekends from the last weekend of 2007 to the last weekend of 2017. To this end, we need to prepare the sample forecast return distributions, which we use to analyze the attributions. In a practical analysis, we analyze the forecast return distribution which we use to construct the optimal portfolio at the beginning of the investment period. Here, we evaluate a test distribution instead of an arbitrary distribution. Remember that a return distribution has a return vector and a covariance matrix. First, we use the realized covariance matrix of the sectoral returns for the preceding 52 weeks until a weekend as the test covariance matrix. Next, we use the realized return vector of the next 52 weeks with some random noises as the test return vector for the weekend. We calculate the next 52 weeks' 10-dimensional test forecast return vector $\boldsymbol{r_F}(t)$ and the 10×10 test covariance matrix $\boldsymbol{\Sigma_F}(t)$ for the next 52 weeks for weekend t as:

$$\boldsymbol{r_F}(t) \equiv \boldsymbol{r_R}(t+52) + c\boldsymbol{s}(t)\boldsymbol{\varepsilon}, \tag{4.3.1}$$

$$\boldsymbol{\Sigma_F}(t) \equiv \boldsymbol{\Sigma_R}(t), \tag{4.3.2}$$

where $\boldsymbol{r_R}(t)$ is the 52 weeks' 10-dimensional realized return vector until weekend t, $\boldsymbol{\Sigma_R}(t)$ is the 10×10 realized covariance matrix of the sectoral returns for the preceding 52 weeks until weekend t,

$$\boldsymbol{s}(t) \equiv \begin{pmatrix} s_1(t) & & 0 \\ & \ddots & \\ 0 & & s_{10}(t) \end{pmatrix}, \boldsymbol{\varepsilon} \equiv \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_{10} \end{pmatrix}, \tag{4.3.3}$$

where $s_i(t)$ are the standard deviations of each sector's return for the preceding 52 weeks until weekend t, $\boldsymbol{z}_i$ are the random variables under $\boldsymbol{N}(0,1)$, each pair $\boldsymbol{z}_i$ and $\boldsymbol{z}_j$ is independent, and c is a scalar coefficient. In our numerical examples, c represents the predictability level of forecast returns.

We calculate the test forecast return vectors, test covariance matrices, realized return vectors, and realized covariance matrices from the first weekend of 2009 to the last weekend of 2016 in cases of $c = 0.1, 0.3,$ or 0.5. The random variables are given from the random function of Excel in each case. The diagonalizations of the covariance matrices are performed using "eigen" function of R (R Core Team. (2016)). From the diagonalizations, we consider the eigen vectors as the risk factor vectors and the eigen values as the risk of the factors. Furthermore, we perform the attribution analysis for the test forecast returns, forecast magnitude of test risks, and test risk factors using Property 4.1 in each case. Table 4.1 summarizes the attribution analysis. The first row shows the average for the entire period, from the first weekend of 2009 to the last weekend of 2016. The other rows show the averages for each year, including the average values of the 53 weeks in 2009 and 2015. The remaining years show the average values of the 52 weeks for those years.

Table 4.1: Attribution of the Sharpe ratio
to the predictability-based measurement of the forecast return and risk

(A) Year	(B) Realized Sharpe Ratio			(C) Marke Condition	(D) The Predictability-Based Measurement of Forecast												
								(E) Forecast Rerutn			(F) Forecast Risk	(G) Magni -tude Fore-cast	(H) Factor Fore-cast	(X) Dupli-cate Term X	(Y) Duplicate Term Y		
	c = 0.1	0.3	0.5		c = 0.1	0.3	0.5	c = 0.1	0.3	0.5					c = 0.1	0.3	0.5
2009-2016	0.276	0.163	0.103	0.482	0.567	0.331	0.210	0.743	0.387	0.255	0.750	0.963	0.759	-0.027	-0.075	-0.195	-0.206
2016	0.303	0.176	0.111	0.543	0.552	0.317	0.198	0.867	0.496	0.305	0.655	0.960	0.716	0.021	-0.030	-0.165	-0.238
2015	0.213	0.106	0.072	0.347	0.616	0.307	0.213	0.804	0.427	0.303	0.776	0.980	0.796	0.001	-0.036	-0.104	-0.134
2014	0.276	0.211	0.176	0.434	0.635	0.477	0.403	0.832	0.439	0.311	0.688	0.965	0.675	-0.048	-0.115	-0.350	-0.404
2013	0.391	0.236	0.079	0.545	0.717	0.429	0.145	0.824	0.388	0.160	0.805	0.964	0.811	-0.030	-0.088	-0.236	-0.180
2012	0.405	0.242	0.166	0.572	0.703	0.414	0.268	0.775	0.420	0.294	0.791	0.943	0.834	-0.013	-0.138	-0.204	-0.183
2011	0.259	0.187	0.116	0.455	0.572	0.408	0.252	0.710	0.401	0.260	0.720	0.974	0.692	-0.055	-0.141	-0.287	-0.272
2010	0.180	0.067	0.038	0.503	0.353	0.127	0.073	0.456	0.206	0.124	0.807	0.968	0.765	-0.073	-0.090	-0.114	-0.142
2009	0.185	0.081	0.064	0.461	0.394	0.176	0.131	0.673	0.316	0.278	0.755	0.951	0.783	-0.021	0.033	-0.105	-0.098

Data Source: Ten sector total indices. S&P Dow Jones Indices. 2018. S&P Dow Jones Indices LLC.

In the first row in Table 1, Column (B) shows the realized Sharpe ratio $SR(\boldsymbol{r_F}, \boldsymbol{\Sigma_F} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R})$ of (4.2.1), which the optimal portfolio based on the test forecast distribution achieves. The optimal portfolios based on the test distribution achieve 0.276 of the realized Sharpe ratio, on average, for the entire period when $\mathrm{c} = 0.1$. Column (C) shows that the Sharpe ratio for the market condition is 0.482, on average, indicating that the maximum Sharpe ratio an investor can obtain is 0.482 through a non-leveraged long/short portfolio. Column (D) shows that the predictability-based measurement of the test distribution is 0.567 on average when $\mathrm{c} = 0.1$. Further the imperfect information on the realized distribution declines by approximately forty-three percent of the achievable maximum Sharpe ratio, on average. Column (B) shows the natural outcome that the optimal portfolio tends to achieve a higher Sharpe ratio when the predictability-based measurement of the forecast is higher in Column (D). Column (E) shows the predictability-based measurement of the test return $P\_SR(\boldsymbol{r_F}, \boldsymbol{\sigma_R}, \boldsymbol{K_R})$ of (4.2.4). Column (F) shows the predictability-based measurement of the test risk $P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})$. Indeed, the predictability-based measurement of test distribution in Column (D) is attributed to the predictability-based measurement of test returns and risks without the duplicate term Y in Column (Y) from (4.2.9). The predictability-based measurement of the test return is 0.743 and that the predictability-based measurement of test risk is 0.750 when $\mathrm{c} = 0.1$. As such, the forty-three percent reduction is, on average, attributed to the twenty-six percent reduction caused by imperfect information of the realized return, and to the twenty-five percent reduction caused by imperfect information of the realized risk. Consequently, we can compare the two predictability-based measurements, simultaneously. The magnitude of the duplicate is 7.5 percent when $\mathrm{c} = 0.1$ from Column (Y). Column (G) shows the predictability-based measurement of the magnitude of the test risk $P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_R})$ of (4.2.5). Column (H) shows the predictability of the test risk factors $P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_F})$ of (4.2.6). The predictability-based measurement of the test risk in Column (F) is attributed to the predictability-based measurement of risk magnitude and risk factors without the duplicate term X in Column (X) from (4.2.10). We find that the predictability-based measurement of magnitude of test risk is 0.963 and that of the risk factors is 0.759. Similarly, the twenty-five percent reduction caused by imperfect information on the realized risk is, on average, attributed to the four percent reduction caused by imperfect information on the magnitude of realized risk, and the twenty-four percent reduction caused by imperfect information on the realized risk factors. As a result, we can identify the main source of the reduction in the predictability-based measurement of the risk. Column (X) shows that the magnitude of the duplicate is 2.7 percent. The predictability-based measurement of test risk factors causes most of the predictability-based measurement of test risk in Columns (F), (G), and (H) of the historical data. From the attribution analyses, we can thus comparatively identify the main source of the reduction in the realized Sharpe ratio. This information may be useful for improving the next investments.

When comparing the rows of years 2016 and 2013, the predictability-based measurements of the test return in 2016 is, on average, higher than in 2013 when $c = 0.1$ and $c = 0.3$. However, the realized Sharpe ratios in 2016 are, on average, lower than those in 2013, although the market conditions in both years have the same level. We identify the reason for the difference in the predictability-based measurement of the test risks between 2016 and 2013 in Column (F). The average predictability-based measurement of the test risk is 0.655 in 2016 and 0.805 in 2013. This difference comes from the difference between the predictability-based measurement of test risk factors in 2016 and 2013. This shows that the risk forecast is important, as is return forecast. In particular, we realize that forecasting the risk factors is important.

4.4 An attribution analysis against the benchmark

In this section, we discuss the attribution analysis against a given benchmark using the above predictability-based measurement. In the usual ways of analyzing against a benchmark, we do not evaluate the performance of the benchmark. Using the attribution analysis above, we can simultaneously evaluate the ex-ante forecasts and the benchmark performance. For example, we identify two explanations for the Sharpe ratio of the optimal portfolio being higher than that of the benchmark over the investment period: one is that the performance of the optimal portfolio is superior and the other is that the performance of the benchmark is inferior. Indeed, we can assess the explanation that is more acceptable in this situation using attribution analysis.

From (4.2.1), the ex-post realized Sharpe ratio achieved by the optimal portfolio from the ex-ante forecast return distribution $\boldsymbol{N}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F})$ under the ex-post realized return distribution $\boldsymbol{N}(\boldsymbol{r_R}, \boldsymbol{\Sigma_R})$ is $SR(\boldsymbol{r_F}, \boldsymbol{\Sigma_F} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R})$. From Property 4.1, we have three predictions to which the Sharpe ratio can be attributed the predictability-based measurement of the ex-ante forecast returns $P\_SR(\boldsymbol{r_F}, \boldsymbol{\sigma_R}, \boldsymbol{K_R})$, that of the ex-ante forecast magnitude of the risk $P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_R})$, and that of the ex-ante forecast risk factors $P\_SR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_F})$.

Next, we discuss the realized Sharpe ratio of a benchmark. Assume we have an n-dimensional vector of the portfolio of benchmark $\boldsymbol{w_{BM}}$ at the beginning of the investment period. Additionally, suppose that we construct the benchmark portfolio $\boldsymbol{w_{BM}}$ based on an assumed n-dimensional vector of ex-ante forecast returns $\boldsymbol{r_{BM}}$ and an assumed $n \times n$ covariance matrix of the ex-ante forecast $\boldsymbol{\Sigma_{BM}}$ using the mean-variance approach at the beginning of the investment period. After the investment period, assume that the benchmark portfolio gains the realized Sharpe ratio SR_{BM} over the investment period.

From Property 3.2, we have

$$SR_{BM} = SR(\boldsymbol{r_{BM}}, \boldsymbol{\Sigma_{BM}} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R}). \tag{4.4.1}$$

In fact, we have several pairs $(\boldsymbol{r_{BM}}, \boldsymbol{\Sigma_{BM}})$ that hold for the above equation. We measure the performance of the benchmark after the investment period. Therefore, in our attribution analysis, we assume that we construct the benchmark portfolio $\boldsymbol{w_{BM}}$ based on an n-dimensional vector of ex-ante implied forecast returns $\boldsymbol{r_{BM\_I}}$ and $n \times n$ covariance matrix $\boldsymbol{\Sigma_R}$, which is covariance matrix of the ex-post realized returns of the risky assets over the investment period. Indeed, this assumption indicates that we have the following implied forecast return for constructing the benchmark portfolio at the beginning of the investment period (see Appendix 2.3)

$$\boldsymbol{r_{BM\_I}} = \boldsymbol{\Sigma_R w_{BM}}. \tag{4.4.2}$$

From (4.2.4), we have the predictability-based measurement of the implied forecast return of the benchmark:

$$P\_SR(\boldsymbol{r_{BM_I}}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) \equiv \left(\frac{\boldsymbol{\sigma_R}^{-1} \boldsymbol{K_R^{-1}} \boldsymbol{r_{BM_I}}}{\|\boldsymbol{\sigma_R}^{-1} \boldsymbol{K_R^{-1}} \boldsymbol{r_{BM_I}}\|}, \boldsymbol{g} \right). \tag{4.4.3}$$

By comparing $P\_SR(\boldsymbol{r_F}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})$ and $P\_SR(\boldsymbol{r_{BM\_I}}, \boldsymbol{\sigma_R}, \boldsymbol{K_R})$, we can evaluate the ex-ante forecasts that construct the optimal portfolio and evaluate the benchmark performance.

Next, we demonstrate the attribution analysis against the benchmark using the historical data sample, that is the weekly S&P 500 10 sector total return indices. It is natural to use the market weight portfolio as the benchmark. For our purpose, we prepare the sample portfolio, which is the optimal portfolio constructed by a given return and risk forecast. When we use the mean-variance approach to analyze the performance of the portfolio that we construct based on a given forecast return and risk, we directly use the forecast return and risk. However, in this section, we use test returns and risk instead of an arbitrary distribution. Here, we compare the performances of the equal weight portfolio and of the benchmark, which is the performance achieved by the market weight portfolio. Additionally, we use the test risk, as in the previous section, to construct the equal weight portfolio using the mean-variance approach. The test risk is the realized covariance matrix of the sectoral returns for the preceding 52 weeks of a weekend. Then, we also have the implied forecast return $\boldsymbol{r_{Eq\_I}}$, which leads to an equal weight portfolio serving as the optimal portfolio when using the test risk and the mean-variance approach.

$$\boldsymbol{r_{Eq\_I}} = \boldsymbol{\Sigma_F w_{Eq}} \,, \tag{4.4.4}$$

where $\boldsymbol{w_{Eq}} = (0.1, \cdots, 0.1)'$ because we have the weekly S&P 500 10 sector total return indices as investable assets in this numerical analysis

We perform the attribution analysis when comparing the predictability-based measurement of the implied return of equal weight portfolio with test risk and the predictability-based measurement of the benchmark.

Table 4.2: Sharpe ratio attribution against the benchmark

(A)	(B)			(C)	(D)										
	Realized Sharpe Ratio			Market	The Predictability-based measurement										
Week End of Year	during the next year			Condition				(E) Rerutn			(F) Risk				(Y)
	Portfolio				Portfolio			Portfolio				(G)	(H)	(X)	
	Equaly Weight-ed	Bench Mark	Excess		Equaly Weight-ed	Bench Mark	Excess	Equaly Weight-ed	Bench Mark	Excess		Magni-tude Forecast	Factor Forecast	Dupli-cate Term X	Dupli-cate Term Y
2016	0.402	0.483	-0.081	0.562	0.716	0.859	-0.143	0.860	0.859	0.001	0.511	0.979	0.707	0.175	-0.345
2015	0.169	0.129	0.040	0.416	0.407	0.311	0.096	0.519	0.311	0.208	0.841	0.987	0.847	-0.007	-0.047
2014	-0.011	0.010	-0.022	0.400	-0.029	0.026	-0.055	-0.268	0.026	-0.294	0.787	0.970	0.820	0.002	-0.452
2013	0.175	0.177	-0.001	0.597	0.294	0.296	-0.002	0.415	0.296	0.119	0.666	0.881	0.642	-0.142	-0.213
2012	0.378	0.407	-0.029	0.675	0.560	0.603	-0.043	0.275	0.603	-0.328	0.860	0.985	0.839	-0.037	-0.425
2011	0.155	0.154	0.000	0.363	0.427	0.426	0.001	0.356	0.426	-0.070	0.813	0.943	0.835	-0.035	-0.258
2010	0.022	0.011	0.012	0.504	0.044	0.021	0.023	0.119	0.021	0.098	0.690	0.976	0.581	-0.134	-0.235
2009	0.123	0.112	0.010	0.482	0.255	0.233	0.021	-0.037	0.233	-0.270	0.802	0.990	0.765	-0.048	-0.489

Data Source: Ten sector total indices. S&P Dow Jones Indices. 2018. S&P Dow Jones Indices LLC.

(Y): Duplicate Term Y of the equaly weighted portfolio

Table 4.2 summarizes the Sharpe ratio attribution analysis against the benchmark for equal weight portfolios for the last weekends of the eight years from 2009 to 2016. In the first row in Table 4.2, Column (B) indicates that the equal weight portfolio for the last weekend of 2016 is 0.402 of the realized Sharpe ratio for 2017. Additionally, the benchmark portfolio, which is the market weight portfolio at the end of 2016, is 0.483 of the realized Sharpe ratio for 2017. The excess Sharpe ratio of the equal weight portfolio is –0.081 against the benchmark. In other words, the Sharpe ratio of the portfolio is inferior to the Sharpe ratio of the benchmark. Column (D) provides the predictability-based measurement of the portfolio and the benchmark. The predictability-based measurement of the portfolio is 0.716 and that of the benchmark is 0.859. The inferiority of the predictability-based measurement of the portfolio leads the inferiority of the realized Sharpe ratio of the portfolio. Column (E) provides the return predictability-based measurement of the portfolio and the benchmark. Both predictability-based measurements are at the same level. Note that the returns are implied returns, indicating that the equal weight portfolio and the benchmark portfolio are the optimal portfolios when using the mean-variance approach. We illustrate the predictability-based measurement of the portfolio risk and its decomposition in Columns (F), (G), and (H). Note that the predictability-based measurement of the benchmark risk is one because of the definition of the implied returns of the benchmark portfolio. Therefore, we realize that the inferiority results from the lower predictability-based measurement of the risk, especially the lower predictability-based measurement of the risk factor.

Additionally, when comparing the portfolio of the last weekend of 2016 to the portfolio of the last weekend of 2015. The equal weight portfolio of the last weekend of 2015 is 0.169 of the realized Sharpe ratio for 2016. The predictability-based measurement of the portfolio is 0.407, which indicates that the predictability-based measurement of the portfolio for the end of 2015 is inferior to that for 2016. Additionally, the predictability of the benchmark for the end of 2015 is 0.311. Therefore, the Sharpe ratio of the portfolio is superior to the Sharpe ratio of the benchmark for 2016. Further, the Sharpe ratio of the portfolio is inferior to the benchmark for 2017, although the predictability-based measurement of the portfolio at the end of 2016 is better than that of the portfolio at the end of 2015. We recognize that the predictability-based measurement of the benchmark at the end of 2016 is superior to that of the portfolio at the end of 2016.

We also determine the importance of the predictability-based measurement of risk when we observe the portfolio at the end of 2013 and its benchmark. From Column (E), the predictability-based measurement of the portfolio return is 0.415, which is superior to the 0.296 of the benchmark return. However, the predictability-based measurement of the portfolio is 0.294, which is close to the 0.296 of the benchmark from Column (D). The predictability-based measurement of the risk of the portfolio is 0.666, this inferior predictability of risk reducing the superior return predictability. We thus obtain rich information from the attribution analysis.

4.5 Summary

In this chapter, we propose an attribution method of the ex-post realized Sharpe ratio achieved by the optimal portfolio based on the ex-ante forecast return distribution. Furthermore, we decompose the realized Sharpe ratio into the absolute value of the market Sharpe ratio, return predictability, risk predictability, which is in turn decomposed as the predictability of the magnitude of risk and predictability of risk factor. Additionally, we present numerical examples of the attribution analysis.

We find that the attribution formula quantitatively enables a comparative analysis between return and risk predictabilities. Furthermore, the predictability of the risk has a more detailed predictability, as the magnitudes of the risk and risk factors. Additionally, we simultaneously evaluate the quality of the investor's portfolio and of the benchmark against the benchmark. These attribution analyses provide rich information for investment analysis, thus being a useful analytical tool for portfolio managers and sponsors.

Chapter 5

Implication of optimal confidence level in Black-Litterman model

5.1 Black-Litterman model

Investors face problems when they make their portfolios reflect their forecasts using the mean-variance approach from modern portfolio theory. One of the problems is that a resulting portfolio is too sensitive to the input parameters. A second problem arises when a portfolio has few assets. Black and Litterman (1992) solve these two problems by merging the confident forecasts with the market consensus forecasts from the capital asset pricing model. Their approach has a key parameter of the confidence level for the investor's confident forecasts.

Fabozzi, et al. (2006) state, "In practice, the Black-Litterman model has proven successful in mitigating estimation errors in the mean-variance framework. The Black-Litterman model does not require that the portfolio manager provide forecasts on all the securities in the universe managed. Rather, it 'blends' any views (this could be a forecast on just one or a few securities, or all them) the investor might have with market equilibrium," and, "When no views are present, the resulting Black-Litterman expected returns are just the expected returns consistent with market equilibrium. Conversely, when the investor has views on some of the asset, the resulting expected return deviate from market equilibrium."

In the Black-Litterman model, a key parameter is introduced and is a parameter that expresses the confidence level of the investor's view. The confidence level of the investor's view indicates the relative importance of the equilibrium versus the views. The confidence level varies with the predictability-based measurement of expected returns, and the Black-Litterman model blends the investor's views and returns that are consistent with the market equilibrium. A numerical sample of this phenomenon is provided in this chapter. An implication of the phenomenon becomes a reference matter when considering the confidence level of the Black-Litterman model.

For active investments in which investors seek returns that are superior to the market's average returns, investors reflect in their investment decisions their forecasts for the returns and risk of investable assets.

In this chapter, we discuss the confidence level of an investor's forecast in the Black-Litterman model, which merges his or her forecasts with the confidence of some risky assets and the consensus forecasts of a number of risky assets with the confidence level parameter. However, the confidence level parameter affects the weight balance between information on the confident forecasts and information on the consensus forecasts in the merged forecasts. Because the number of risky assets that the investor provides to the confident forecast is relatively small in general, overconfidence may affect the proportion of the forecasted Sharpe ratios of the forecast factors in the merged forecast. We find that overconfidence misses opportunities to obtain a higher Sharpe ratio despite the precise forecast in the sample case.

First, we provide a brief review of the Black-Litterman model. Let $\boldsymbol{\mu_F}$ be the forecast returns of n risky assets. Additionally, let the investor's forecast with confidence be the following

$$\boldsymbol{P\mu_F} \sim \boldsymbol{N}(\boldsymbol{v}, \boldsymbol{\Omega}), \tag{5.1.1}$$

where $\boldsymbol{P}$ is a pickup matrix that selects the factors that are forecast with confidence, $\boldsymbol{v}$ is the forecast return, and $\boldsymbol{\Omega}$ is the confidence level. For example, if an investor has confidently forecasted that the return of the first risky asset is five percent, then $\boldsymbol{P} = (1, 0, \cdots, 0), \boldsymbol{v} = 5\%$, and $\boldsymbol{\Omega} = \omega$. In this case, $\boldsymbol{v}$ is a scalar and $\boldsymbol{\Omega}$ is a 1×1 matrix because the investor only has a confident forecast for one asset. The model refers to the market consensus returns and risks as another forecast from the CAPM. Therefore, the market consensus return $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_n)'$ is derived from the market risk matrix $\boldsymbol{\Sigma_M}$ and the current market portfolio $\boldsymbol{W_M}$ as in the following

$$\boldsymbol{\pi} = 2\theta \boldsymbol{\Sigma_M W_M}, \tag{5.1.2}$$

where θ is the risk premium. As we merge the investor's forecast with the market consensus using the Black-Litterman model, we obtain the merged forecast of the returns $\boldsymbol{\mu_{BL}(v)}$ and the risk $\boldsymbol{\Sigma_{BL}}$

$$\boldsymbol{\mu_{BL}(v)} = \boldsymbol{\pi} + \boldsymbol{\Sigma_M P'(P\Sigma_M P' + \Omega)^{-1}(v - P\pi)}, \tag{5.1.3}$$

$$\boldsymbol{\Sigma_{BL}} = \boldsymbol{\Sigma_M} - \boldsymbol{\Sigma_M P'(P\Sigma_M P' + \Omega)^{-1} P\Sigma_M}. \tag{5.1.4}$$

5.2 Implication of optimal confidence level

We remember expression (4.2.4) of the Sharpe ratio, which expresses the ex-post realized Sharpe ratio that the optimal portfolio from the ex-ante forecast return distribution $\boldsymbol{N(r_F, \Sigma_F)}$ achieves under the ex-post realized return distribution $\boldsymbol{N(r_R, \Sigma_R)}$.

$$SR(\boldsymbol{r_F}, \boldsymbol{\Sigma_F} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R}) = \left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\| \left(\frac{\boldsymbol{\Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}}{\left\|\boldsymbol{\Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}\right\|}, \boldsymbol{g} \right), \tag{5.2.1}$$

where

$$\boldsymbol{\Phi} \equiv \boldsymbol{\sigma_R K_R^{-1} K_F \sigma_F}^{-1}, \boldsymbol{g} \equiv \frac{\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}}{\left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\|}, \tag{5.2.2}$$

and $(\boldsymbol{x}, \boldsymbol{y})$ is the inner product of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$.

Therefore, from (5.1.3) and (5.1.4), the optimal portfolio based on the merged forecast returns and risk using the Black-Litterman model achieves the Sharpe ratio under the ex-post realized return distribution $\boldsymbol{N(r_R, \Sigma_R)}$

$$SR(\boldsymbol{\mu_{BL}(v)}, \boldsymbol{\Sigma_{BL}} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R}), \tag{5.2.3}$$

under the ex-post realized return distribution $\boldsymbol{N(r_R, \Sigma_R)}$. Supposing that $\boldsymbol{K_{BL}^{-1} \Sigma_{BL} K_{BL}} = \boldsymbol{\sigma_{BL}^2}$, we have

$$SR(\boldsymbol{\mu_{BL}(v)}, \boldsymbol{\Sigma_{BL}} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R}) = \left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\| \left(\frac{\boldsymbol{\Phi_{BL} \sigma_{BL}}^{-1} \boldsymbol{K_{BL}^{-1} \mu_{BL}(v)}}{\left\|\boldsymbol{\Phi_{BL} \sigma_{BL}}^{-1} \boldsymbol{K_{BL}^{-1} \mu_{BL}(v)}\right\|}, \boldsymbol{g} \right), \tag{5.2.4}$$

where

$$\boldsymbol{\Phi_{BL}} \equiv \boldsymbol{\sigma_R K_R^{-1} K_{BL} \sigma_{BL}}^{-1}, \boldsymbol{g} \equiv \frac{\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}}{\left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\|}. \tag{5.2.5}$$

In contrast, the market consensus return $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_n)'$ is derived from the market risk matrix $\boldsymbol{\Sigma_M}$ and the current market portfolio $\boldsymbol{W_M}$ as in the following from (5.1.2). That the market consensus does not necessarily have the perfect risk forecast for the investable assets is natural. Indeed, we suppose $\boldsymbol{\Sigma_M} = \boldsymbol{\Sigma_F}$. From (5.1.2), we have

$$\boldsymbol{\pi} = 2\theta \boldsymbol{\Sigma_F W_M}. \tag{5.2.6}$$

Therefore, the market portfolio achieves the Sharpe ratio

$$SR(\boldsymbol{\pi}, \boldsymbol{\Sigma_F} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R})$$
$$= \left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\| \left(\frac{\boldsymbol{\Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1} \pi}}{\left\|\boldsymbol{\Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1} \pi}\right\|}, \boldsymbol{g}\right) = \left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\| \left(\frac{\boldsymbol{\sigma_F K_F^{-1} W_M}}{\left\|\boldsymbol{\sigma_F K_F^{-1} W_M}\right\|}, \boldsymbol{g}\right). \tag{5.2.7}$$

From (5.2.4) and (5.2.7), we recognize that the magnitude relation between

$$\left(\frac{\boldsymbol{\Phi_{BL} \sigma_{BL}}^{-1} \boldsymbol{K_{BL}^{-1} \mu_{BL}(v)}}{\left\|\boldsymbol{\Phi_{BL} \sigma_{BL}}^{-1} \boldsymbol{K_{BL}^{-1} \mu_{BL}(v)}\right\|}, \boldsymbol{g}\right) \text{ and } \left(\frac{\boldsymbol{\sigma_F K_F^{-1} W_M}}{\left\|\boldsymbol{\sigma_F K_F^{-1} W_M}\right\|}, \boldsymbol{g}\right) \tag{5.2.8}$$

determines whether the optimal portfolio based on the merged forecast achieves a higher Sharpe ratio than the market portfolio.

We execute a simple case study where the investor has the perfect forecast for the risk factors and directly invests in the factors. Indeed, $\boldsymbol{K_F} = \boldsymbol{K_R} = \boldsymbol{I}$. Additionally, the investor directly forecasts the first factor return with confidence. From (5.1.3) and (5.1.4), we have

$$\boldsymbol{\mu_{BL}(v)} = \boldsymbol{\pi} + \boldsymbol{\sigma_R^2 P'}(\boldsymbol{P \sigma_R^2 P'} + \boldsymbol{\Omega})^{-1}(\boldsymbol{v} - \boldsymbol{P\pi}), \tag{5.2.9}$$
$$\boldsymbol{\Sigma_{BL}} = \boldsymbol{\sigma_R^2} - \boldsymbol{\sigma_R^2 P'}(\boldsymbol{P \sigma_R^2 P'} + \boldsymbol{\Omega})^{-1} \boldsymbol{P \sigma_R^2}, \tag{5.2.10}$$

where $\boldsymbol{P} = (1, 0, \cdots, 0)$ in this case. We have the merged forecast factor Sharpe ratio vector $\boldsymbol{\sigma_{BL}}^{-1} \boldsymbol{K_{BL}^{-1} \mu_{BL}(v)}$ (see Appendix 2.5)

$$\boldsymbol{\sigma_{BL}}^{-1} \boldsymbol{K_{BL}^{-1} \mu_{BL}(v)} = \left(\frac{\omega^2 \pi_1 + \sigma_R^2(1) v}{\sigma_R(1)\omega\sqrt{\sigma_T^2(1) + \omega^2}}, \frac{\pi_2}{\sigma_R(2)}, \cdots, \frac{\pi_n}{\sigma_R(n)}\right)'. \tag{5.2.11}$$

We rewrite

$$\frac{\omega^2 \pi_1 + \sigma_R^2(1) v}{\sigma_R(1)\omega\sqrt{\sigma_R^2(1) + \omega^2}} = \frac{\pi_1 + \frac{\sigma_R(1)}{\omega^2} v}{\sigma_R(1)\sqrt{\frac{\sigma_R^2(1)}{\omega^2} + 1}} = \frac{\frac{\omega}{\sigma_R(1)}\pi_1 + \frac{\sigma_R(1)}{\omega} v}{\sqrt{\sigma_R^2(1) + \omega^2}}. \tag{5.2.12}$$

As $\omega \to 0$, this value $\to \infty$ because of the term on the far right, even if v is the just precise return $r_R(1)$. Therefore, $\omega \to 0$ and

$$\frac{\boldsymbol{\Phi_{BL} \sigma_{BL}}^{-1} \boldsymbol{K_{BL}^{-1} \mu_{BL}(v)}}{\left\|\boldsymbol{\Phi_{BL} \sigma_{BL}}^{-1} \boldsymbol{K_{BL}^{-1} \mu_{BL}(v)}\right\|} \to (1,0,\cdots,0)'. \tag{5.2.13}$$

In contrast, as $\omega \to \infty$, then

$$\frac{\omega^2 \pi_1 + \sigma_R^2(1) v}{\sigma_R(1)\omega\sqrt{\sigma_R^2(1) + \omega^2}} \to \pi_1/\sigma_R(1) \tag{5.2.15}$$

as the Sharpe ratio of the first factor included the market consensus. Therefore,

$$\boldsymbol{\Phi_{BL} \sigma_{BL}}^{-1} \boldsymbol{K_{BL}^{-1} \mu_{BL}(v)} \to \boldsymbol{\sigma_F K_F^{-1} W_M}\,. \tag{5.2.16}$$

We note that cases exist in which the value of the inner product

$$\left((1,0,\cdots,0)', \frac{\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}}{\left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\|}\right) \tag{5.2.17}$$

is less than the value of the inner product

$$\left(\frac{\boldsymbol{\sigma_F K_F^{-1} W_M}}{\left\|\boldsymbol{\sigma_F K_F^{-1} W_M}\right\|}, \frac{\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}}{\left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\|}\right). \tag{5.2.18}$$

These cases indicate that investor overconfidence may cause the opportunity to achieve a higher Sharpe ratio to be missed when using the Black-Litterman model.

5.3 A numerical example

In this section, we have a simple example. Suppose that only four factors are investable assets and let the ex-post realized distribution of returns be as in Table 5.1.

Table 5.1 Ex-post realized distribution

	Factor 1	Factor 2	Factor 3	Factor 4
Return	8.00%	5.00%	3.00%	1.00%
Risk	15.00%	10.00%	8.00%	6.00%
Sharpe ratio	0.533	0.500	0.375	0.167

The optimal portfolio based on the true distribution is $(22.2\%, 31.2\%, 29.3\%, 17.3\%)'$. The percentages in the parentheses are the weights of the factors as investable assets. The optimal portfolio achieves the maximum Sharpe ratio of 0.84. We suppose that we have the perfect risks for all factors. However, we suppose that the market consensus forecast has an imperfect return forecast. Then, we suppose that the return forecast of Factor 1 is 2.00%, the return forecast of Factor 2 is 5.96%, the return forecast of Factor 3 is 3.58%, and the return forecast of Factor 4 is 1.19%, respectively. Of course, these numbers are artificial. The optimal portfolio based on the market consensus forecasts using the mean-variance approach is then $(1.47\%, 39.5\%, 37.1\%, 22.0\%)'$. The optimal portfolio has a Sharpe ratio of 0.67 under the ex-post realized distribution, which does not reach the maximum Sharpe ratio. We suppose that an investor has a perfect forecast of the return to Factor 1 and that the investor has strong confidence in the forecast return of Factor 1 of 8.0%. However, the investor has no forecast for the other factors and, thus, uses the Black-Litterman model with the market consensus forecast. The Sharpe ratio of the portfolio converges to 0.533 as the Sharpe ratio of Factor 1 when $\omega \to 0$. In contrast, suppose that the investor does not have strong confidence. For example, if the investor has low confidence that $\omega = 0.291$, then the investor has the best portfolio $(22.2\%, 31.2\%, 29.3\%, 17.3\%)'$. Moreover, the investor's optimal merged portfolio using the Black-Litterman model achieves the maximum Sharpe ratio.

5.4 Summary

The Black-Litterman model has a key parameter, which is the confidence level of the investor's view. This confidence level indicates the relative importance of the equilibrium versus the views and varies the proposed predictability-based measurement of the expected returns from the Black-Litterman model. A numerical example is provided in which strong investor confidence reduces the realized Sharpe ratio more than the maximum Sharpe ratio.

An implication of the phenomenon indicates an issue in considering the confidence level of the Black-Litterman model.

Chapter 6

Predictability-based measurement for realized information ratio and its attribution to the ex-ante forecast

6.1 Introduction

Along with the Sharpe ratio, the information ratio is another important measure of investment performance. In this chapter, we discuss the information ratio. First, Theorem 6.1 is provided, which addresses the returns that provide an optimal portfolio under an ex-post realized risk that achieves the information ratio identical to that achieved by the optimal portfolio based on ex-ante forecast returns and risk.

Next, a measurement of the power of an ex-ante forecast that predicts the ex-post realized information ratio is proposed, similar to the case of the Sharpe ratio. In this thesis, we call the measurement *the predictability-based measurement for the information ratio*. Similar to the Sharpe ratio, we find that the ex-post realized information ratio is attributed to the realized market conditions, the predictability-based measurement of the returns, the predictability-based measurement of the magnitude of the risk, and the predictability-based measurement of the risk factors. Note that these predictability-based measurements are for the information ratio. Additionally, we provide numerical examples of the information ratio attribution analysis and find that the forecast risk effect differs between the reduction in the Sharpe ratio from the achievable Sharpe ratio and the reduction in the information ratio from the achievable information ratio.

6.2 Formulation

In this chapter, we discuss the information ratio. The active portfolio, which is defined as the excess weight vector against the weight vector of the benchmark portfolio, plays an important role in the discussion. If an investor has a portfolio $\boldsymbol{w}$, his or her active portfolio is

$$\boldsymbol{w_A} = \boldsymbol{w} - \boldsymbol{w_{BM}}. \tag{6.2.1}$$

Note that the sum of the active weights of each risk asset $\boldsymbol{e}'\boldsymbol{w_A}$ is zero.

Investors have active portfolios because they believe that they have richer information on n risky assets against the benchmark in the future investment period. In the context of the mean-variance approach, the richer information is the forecast return distribution itself. We assume that the investor has an ex-ante forecast with an n-dimensional normal distribution $\boldsymbol{N}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F})$ at the beginning of the investment period. Suppose we have a benchmark portfolio $\boldsymbol{w_{BM}}$ at the beginning of the investment period. The investor expects an excess return against the benchmark

$$\boldsymbol{r_F}'\boldsymbol{w} - \boldsymbol{r_F}'\boldsymbol{w_{BM}} = \boldsymbol{r_F}'\boldsymbol{w_A}. \tag{6.2.2}$$

The information ratio is defined as the value of the excess return divided by the risk of the excess return. The risk is measured by the standard deviation of the excess return. Therefore, the investor constructs the optimal portfolio that achieves the maximum information ratio by solving the following quadratic programming problem in the context of the mean-variance approach.

For a given number h, we find an optimal active portfolio $\boldsymbol{w_F}$ against the benchmark portfolio $\boldsymbol{w_{BM}}$ as the solution to the following quadratic programming problem:

$$minimize\ \boldsymbol{w_F}'\boldsymbol{\Sigma_F}\boldsymbol{w_F},$$

$$\text{s.t. } \boldsymbol{r_F}'\boldsymbol{w_F} = h, \text{and } \boldsymbol{e}'\boldsymbol{w_F} = 0, \tag{6.2.3}$$

where $\boldsymbol{e} = (1, \cdots, 1)'$.

Let c be an arbitrary number. Problem (6.2.3) has the following solution (see Appendix 2.6)

$$\boldsymbol{w_F}(c) = c\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}(\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e}), \tag{6.2.4}$$

where

$$\alpha_F \equiv \boldsymbol{e}'\boldsymbol{\Sigma_F}^{-1}\boldsymbol{e}, \beta_F \equiv \boldsymbol{e}'\boldsymbol{\Sigma_F}^{-1}. \tag{6.2.5}$$

Note that the active portfolio $\boldsymbol{w_F}(c)$ achieves the same maximum information ratio for an arbitrary number c.

We have the following theorem between the ex-ante forecast distribution and the ex-post realized distribution.

Theorem 6.1:

The optimal active portfolio based on the return distribution $\boldsymbol{N}(\boldsymbol{r}_{Trans}, \boldsymbol{\Sigma_R})$ according to the information ratio, where the two parameters of the return distribution $\boldsymbol{N}(\boldsymbol{r}_{Trans}, \boldsymbol{\Sigma_R})$ are a return $\boldsymbol{r}_{Trans} = \frac{\delta_R}{\alpha_R \delta_F} \boldsymbol{K_R \sigma_R \Phi \sigma_F}^{-1}\boldsymbol{K}_{\boldsymbol{F}}^{\boldsymbol{-1}}(\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e}) + \frac{\beta_R}{\alpha_R}\boldsymbol{e}$ and the ex-post realized covariance matrix $\boldsymbol{\Sigma_R}$, achieves an information ratio identical to the information ratio of the optimal active portfolio based on the ex-ante forecast return distribution $\boldsymbol{N}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F})$, where

$$\alpha_F \equiv \boldsymbol{e}'\boldsymbol{\Sigma_F}^{-1}\boldsymbol{e}, \beta_F \equiv \boldsymbol{e}'\boldsymbol{\Sigma_F}^{-1}\boldsymbol{r_F}, \gamma_F \equiv \boldsymbol{r_F}'\boldsymbol{\Sigma_F}^{-1}\boldsymbol{r_F}, \delta_F \equiv \alpha_F\gamma_F - \beta_F^2,$$
$$\alpha_R \equiv \boldsymbol{e}'\boldsymbol{\Sigma_R}^{-1}\boldsymbol{e}, \beta_R \equiv \boldsymbol{e}'\boldsymbol{\Sigma_R}^{-1}\boldsymbol{r_R}, \gamma_R \equiv \boldsymbol{r}_{\boldsymbol{R}}'\boldsymbol{\Sigma_R}^{-1}\boldsymbol{r_R}, \delta_R \equiv \alpha_R\gamma_R - \beta_R^2, \text{and}$$
$$\boldsymbol{\Phi} \equiv \boldsymbol{\sigma_R K_R^{-1} K_F \sigma_F}^{-1}. \tag{6.2.6}$$

Next, we consider the information ratio $\mathrm{IR}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R})$ that the optimal active portfolio $\boldsymbol{w_F}(c) = c\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}(\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e})$ achieves under the ex-post realized return distribution $\boldsymbol{N}(\boldsymbol{r_R}, \boldsymbol{\Sigma_R})$.

$$\mathrm{IR}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R})$$
$$= \frac{\left(\boldsymbol{\Phi\sigma_F}^{-1}\boldsymbol{K_F^{-1}}(\alpha_F\boldsymbol{r_F} - \beta_F\boldsymbol{e})\right)'(\boldsymbol{\sigma_R}^{-1}\boldsymbol{K_R^{-1}r}_R)}{\sqrt{\left(\boldsymbol{\Phi\sigma_F}^{-1}\boldsymbol{K_F^{-1}}(\alpha_F\boldsymbol{r_F} - \beta_F\boldsymbol{e})\right)'\left(\boldsymbol{\Phi\sigma_F}^{-1}\boldsymbol{K_F^{-1}}(\alpha_F\boldsymbol{r_F} - \beta_F\boldsymbol{e})\right)}}. \tag{6.2.7}$$

We summarize this equation as the following property.

Property 6.1

The optimal active portfolio based on the ex-ante forecast return distribution $\boldsymbol{N}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F})$

achieves the information ratio $\mathrm{IR}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R})$ under the ex-post realized return distribution $\boldsymbol{N}(\boldsymbol{r_R}, \boldsymbol{\Sigma_R})$:

$$\mathrm{IR}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R}) = \left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}\right\|(\boldsymbol{f}, \boldsymbol{g}), \tag{6.2.8}$$

where

$$\boldsymbol{f} \equiv \frac{\boldsymbol{\Phi}\boldsymbol{\sigma_F}^{-1}\boldsymbol{K}_{\boldsymbol{F}}^{-\mathbf{1}}(\alpha_F\boldsymbol{r_F} - \beta_F\boldsymbol{e})}{\left\|\boldsymbol{\Phi}\boldsymbol{\sigma_F}^{-1}\boldsymbol{K}_{\boldsymbol{F}}^{-\mathbf{1}}(\alpha_F\boldsymbol{r_F} - \beta_F\boldsymbol{e})\right\|}, \boldsymbol{g} \equiv \frac{\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}}{\left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}\right\|}, \tag{6.2.9}$$

and $(\boldsymbol{f}, \boldsymbol{g})$ is the inner product of f and $\boldsymbol{g}$. Note that $\boldsymbol{f}$ is the normalized forecast transferred return per risk ratio vector, and that $\boldsymbol{g}$ is the normalized ex-post realized factor Sharpe ratio vector. $\mathrm{SR}(\boldsymbol{r_R}, \boldsymbol{\Sigma_R} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R}) = \left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}\right\|$ from (3.3.11). Therefore, the information ratio does not exceed the maximum Sharpe ratio from (6.2.8).

From (6.2.8) and (6.2.9), we find a different phenomenon of the forecasting power of the information ratio from the forecasting power of the Sharpe ratio. For the Sharpe ratio, the optimal portfolio based on the perfect forecast achieves the highest Sharpe ratio. From (3.3.11), remember that the Sharpe ratio $\mathrm{SR_F}$, which is achieved by the optimal portfolio based on an ex-ante forecast return distribution $\boldsymbol{N}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F})$ under an ex-post realized return distribution $\boldsymbol{N}(\boldsymbol{r_R}, \boldsymbol{\Sigma_R})$, is

$$\mathrm{SR_F} = \left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}\right\|(\boldsymbol{f}, \boldsymbol{g}), \tag{6.2.10}$$

where

$$\boldsymbol{f} \equiv \frac{\boldsymbol{\Phi}\boldsymbol{\sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{K}_{\boldsymbol{F}}^{-\mathbf{1}}\boldsymbol{r_F}}{\left\|\boldsymbol{\Phi}\boldsymbol{\sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{K}_{\boldsymbol{F}}^{-\mathbf{1}}\boldsymbol{r_F}\right\|}, \boldsymbol{g} \equiv \frac{\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}}{\left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}\right\|}, \tag{6.2.11}$$

and $(\boldsymbol{f}, \boldsymbol{g})$ is the inner product of $\boldsymbol{f}$ and $\boldsymbol{g}$. Note that $\boldsymbol{g}$ is the normalized ex-post realized factor Sharpe ratio vector. The perfect forecast indicates that $\boldsymbol{r_F} = \boldsymbol{r_R}$ and $\boldsymbol{\Sigma_F} = \boldsymbol{\Sigma_R}$. Because $\boldsymbol{\Phi} = \mathbf{I}$ in this case, we have $(\boldsymbol{f}, \boldsymbol{g}) = (\boldsymbol{g}, \boldsymbol{g}) = 1$. The perfect forecast provides the highest Sharpe ratio $\left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}\right\|$. In contrast, from (6.2.8), the perfect forecast provides the information ratio

$$\begin{aligned}&\mathrm{IR}(\boldsymbol{r_R}, \boldsymbol{\Sigma_R} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R})\\ &= \left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}\right\|\left(\frac{\boldsymbol{\sigma_R}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}(\alpha_R\boldsymbol{r_R} - \beta_R\boldsymbol{e})}{\left\|\boldsymbol{\sigma_R}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}(\alpha_R\boldsymbol{r_R} - \beta_R\boldsymbol{e})\right\|}, \frac{\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}}{\left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}\right\|}\right).\end{aligned} \tag{6.2.12}$$

The value $\left(\frac{\boldsymbol{\sigma_R}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}(\alpha_R\boldsymbol{r_R} - \beta_R\boldsymbol{e})}{\left\|\boldsymbol{\sigma_R}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}(\alpha_R\boldsymbol{r_R} - \beta_R\boldsymbol{e})\right\|}, \frac{\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}}{\left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}\right\|}\right)$ is not necessarily one. Remember that the value $\left\|\boldsymbol{\sigma}_{\boldsymbol{R}}^{-1}\boldsymbol{K}_{\boldsymbol{R}}^{-\mathbf{1}}\boldsymbol{r_R}\right\|$ is identical to the highest Sharpe ratio that an investor can achieve under a market condition.

6.3 Attribution of information ratio

Recall that the inner product is a measure of the ex-ante forecast returns and risk. We introduce a predictability-based measurement for the information ratio as a quantity $P\_IR(\boldsymbol{r_F}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})$ that has the three parameters $\boldsymbol{r_F}, \boldsymbol{\sigma_F}$, and $\boldsymbol{K_F}$.

$$P\_IR(\boldsymbol{r_F}, \boldsymbol{\sigma_F}, \boldsymbol{K_F}) \equiv \left(\frac{\boldsymbol{\Phi} \boldsymbol{\sigma_F}^{-1} \boldsymbol{K_F^{-1}} (\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e})}{\left\| \boldsymbol{\Phi} \boldsymbol{\sigma_F}^{-1} \boldsymbol{K_F^{-1}} (\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e}) \right\|}, \boldsymbol{g} \right), \tag{6.3.1}$$

as well as

$$P\_IR(\boldsymbol{r_F}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) \equiv \left(\frac{\boldsymbol{\Phi} \boldsymbol{\sigma_R}^{-1} \boldsymbol{K_R^{-1}} (\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e})}{\left\| \boldsymbol{\Phi} \boldsymbol{\sigma_R}^{-1} \boldsymbol{K_F^{-1}} (\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e}) \right\|}, \boldsymbol{g} \right)$$
$$= \left(\frac{\boldsymbol{\sigma_R}^{-1} \boldsymbol{K_R^{-1}} \left(\alpha_R \boldsymbol{r_F} - \boldsymbol{e}' \boldsymbol{\Sigma_R}^{-1} \boldsymbol{r_F} \boldsymbol{e}\right)}{\left\| \boldsymbol{\sigma_R}^{-1} \boldsymbol{K_R^{-1}} \left(\alpha_R \boldsymbol{r_F} - \boldsymbol{e}' \boldsymbol{\Sigma_R}^{-1} \boldsymbol{r_F} \boldsymbol{e}\right) \right\|}, \boldsymbol{g} \right). \tag{6.3.2}$$

It is a predictability-based measurement of a forecast as long as the risk forecast is perfect, $\boldsymbol{\Sigma_F} = \boldsymbol{\Sigma_R}$. Indeed, it means that the forecast of the magnitude of the risk is perfect, $\boldsymbol{\sigma_F} = \boldsymbol{\sigma_R}$, and that the forecast of the risk factors is perfect, $\boldsymbol{K_F} = \boldsymbol{K_R}$. In other words, the predictability-based measurement of the ex-ante forecast returns is not identical to the ex-post realized returns.

Similarly, we have

$$P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_F}) \equiv \left(\frac{\boldsymbol{\Phi} \boldsymbol{\sigma_F}^{-1} \boldsymbol{K_F^{-1}} (\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e})}{\left\| \boldsymbol{\Phi} \boldsymbol{\sigma_F}^{-1} \boldsymbol{K_F^{-1}} (\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e}) \right\|}, \boldsymbol{g} \right)$$
$$= \left(\frac{\boldsymbol{\Phi} \boldsymbol{\sigma_F}^{-1} \boldsymbol{K_F^{-1}} \left(\alpha_F \boldsymbol{r_R} - \boldsymbol{e}' \boldsymbol{\Sigma_F}^{-1} \boldsymbol{r_R} \boldsymbol{e}\right)}{\left\| \boldsymbol{\Phi} \boldsymbol{F}^{-1} \boldsymbol{K_F^{-1}} \left(\alpha_F \boldsymbol{r_R} - \boldsymbol{e}' \boldsymbol{\Sigma_F}^{-1} \boldsymbol{r_R} \boldsymbol{e}\right) \right\|}, \boldsymbol{g} \right). \tag{6.3.3}$$

It is the predictability-based measurement of a forecast as long as the return forecast is perfect.

Similarly, we define the predictability-based measurement of the ex-ante forecast magnitude of the risk that is not identical to the ex-post realized magnitude of the risk

$$P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_R})$$
$$= \left(\frac{\boldsymbol{\sigma_R} \boldsymbol{\sigma_F}^{-2} \boldsymbol{K_R^{-1}} \left(\left(\boldsymbol{e}' \boldsymbol{K_R} \boldsymbol{\sigma_F}^{-2} \boldsymbol{K_R^{-1}} \boldsymbol{e}\right) \boldsymbol{r_R} - \left(\boldsymbol{e}' \boldsymbol{K_R} \boldsymbol{\sigma_F}^{-2} \boldsymbol{K_R^{-1}} \boldsymbol{r_R}\right) \boldsymbol{e} \right)}{\left\| \boldsymbol{\sigma_R} \boldsymbol{\sigma_F}^{-2} \boldsymbol{K_R^{-1}} \left(\left(\boldsymbol{e}' \boldsymbol{K_R} \boldsymbol{\sigma_F}^{-2} \boldsymbol{K_R^{-1}} \boldsymbol{e}\right) \boldsymbol{r_R} - \left(\boldsymbol{e}' \boldsymbol{K_R} \boldsymbol{\sigma_F}^{-2} \boldsymbol{K_R^{-1}} \boldsymbol{r_R}\right) \boldsymbol{e} \right) \right\|}, \boldsymbol{g} \right), \tag{6.3.4}$$

and the predictability-based measurement of the ex-ante forecast risk factors that is not identical to the ex-post realized risk factors

$$P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_F})$$
$$\left(\frac{\boldsymbol{\sigma_R} \boldsymbol{K_R^{-1}} \boldsymbol{K_F} \boldsymbol{\sigma_R}^{-2} \boldsymbol{K_F^{-1}} \left(\left(\boldsymbol{e}' \boldsymbol{K_F} \boldsymbol{\sigma_R}^{-2} \boldsymbol{K_F^{-1}} \boldsymbol{e}\right) \boldsymbol{r_R} - \left(\boldsymbol{e}' \boldsymbol{K_F} \boldsymbol{\sigma_R}^{-2} \boldsymbol{K_F^{-1}} \boldsymbol{r_R}\right) \boldsymbol{e} \right)}{\left\| \boldsymbol{\sigma_R} \boldsymbol{K_R^{-1}} \boldsymbol{K_F} \boldsymbol{\sigma_R}^{-2} \boldsymbol{K_F^{-1}} \left(\left(\boldsymbol{e}' \boldsymbol{K_F} \boldsymbol{\sigma_R}^{-2} \boldsymbol{K_F^{-1}} \boldsymbol{e}\right) \boldsymbol{r_R} - \left(\boldsymbol{e}' \boldsymbol{K_F} \boldsymbol{\sigma_R}^{-2} \boldsymbol{K_F^{-1}} \boldsymbol{r_R}\right) \boldsymbol{e} \right) \right\|}, \boldsymbol{g} \right). \tag{6.3.5}$$

From (6.3.5), $P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) = 1$ does not necessarily hold because

$$P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) = \left(\frac{\boldsymbol{\sigma_R}^{-1}\boldsymbol{K_R^{-1}}(\alpha_R \boldsymbol{r_R} - \beta_R \boldsymbol{e})}{\left\| \boldsymbol{\sigma_R}^{-1}\boldsymbol{K_R^{-1}}(\alpha_R \boldsymbol{r_R} - \beta_R \boldsymbol{e}) \right\|}, \boldsymbol{g} \right). \tag{6.3.6}$$

Again, because the vectors in these inner products are normalized, their values are numbers between one and negative one. Therefore, we recognize the value

$$P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) \tag{6.3.7}$$

as a limit of the information ratio by the perfect forecast. Then, we find that the value

$$P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) - P\_IR(\boldsymbol{r_F}, \boldsymbol{\sigma_F}, \boldsymbol{K_F}) \tag{6.3.8}$$

shows the reduction in the maximum information ratio that occurs because the ex-ante forecast distribution is not identical to the ex-post realized distribution. We have

$$\begin{aligned} &P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) - P\_IR(\boldsymbol{r_F}, \boldsymbol{\sigma_F}, \boldsymbol{K_F}) \\ &= \big(P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) - P\_IR(\boldsymbol{r_F}, \boldsymbol{\sigma_R}, \boldsymbol{K_R})\big) \\ &+\big(P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) - P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})\big) + DTY\_IR, \end{aligned} \tag{6.3.9}$$

where DTY_IR is the duplicate term Y for the information ratio. From (6.3.9), the reduction is attributed to the reduction caused by the fact that ex-ante forecast returns are not identical to ex-post realized returns, the reduction caused by the fact that the ex-ante forecast risk is not identical to the ex-post realized risk, and the duplicate term Y for the information ratio. Both the first and second terms of the right-hand side of (6.3.9) include a synergy reduction caused by the fact that neither the ex-ante forecast returns nor the risk is identical to the ex-post realized returns and risk.

Similarly, we have

$$\begin{aligned} &P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) - P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_F}) \\ &= \big(P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) - P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_R})\big) \\ &+\big(P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) - P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_F})\big) + DTX\_IR, \end{aligned} \tag{6.3.10}$$

where DTX_IR is the duplicate term X. Note that $\big(P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_R}) - P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})\big)$ is the reduction caused by the fact that the ex-ante forecast risk is not identical to the ex-post realized risk. From (6.3.10), the reduction attributes to the reduction caused by the fact that the ex-ante forecast magnitude of the risk is not identical to the ex-post realized magnitude of the risk, the reduction caused by the fact that the ex-ante forecast risk factors are not identical to the ex-post realized risk factors, and the duplicate term X for the information ratio. From (6.3.7), (6.3.8), (6.3.9), and (6.3.10), we finally derive Property 6.2.

Property 6.2:

The ex-post realized information ratio is attributed to seven components. One is the absolute level of the ex-post realized Shape ratio, which represents the market conditions in the investment period. The second is the limit of the information ratio by a perfect forecast. The next three components are the predictability-based measurement of the ex-ante forecast return, the predictability-based measurement of the ex-ante forecast magnitude of risk, and the predictability-based measurement of the ex-ante forecast risk factors, which an investor uses to construct an optimal portfolio with the mean-variance approach at the beginning of the period. The other two are duplicate terms for the information ratio.

6.4 Numerical examples of the attributions

We demonstrate our attribution analysis of the information ratio using the historical data sample from the previous sections. We use ten weekly S&P500 sector total return indices from the first weekend of 2006 to the last weekend of 2017. For our purpose, we need to prepare the sample forecast return distributions, which we use to analyze the attributions. Using a practical analysis, we analyze the forecast return distribution that an investor uses to construct the optimal portfolio at the beginning of the investment period. Here, we evaluate test distributions instead of an investor's arbitrary forecast return distribution. Remember that a return distribution has a return vector and a covariance matrix. First, we use the realized covariance matrix of the sectoral returns for the preceding 52 weeks until a weekend is used as the test covariance matrix. Next, we use the realized return vector of the next 52 weeks with some random noises as the test return vector for the weekend. Indeed, we calculate the next 52 weeks' test return vector $\boldsymbol{r_F}(t)$ and the test covariance matrix $\boldsymbol{\Sigma_F}(t)$ for the next 52 weeks for weekend t as the following:

$$\boldsymbol{r_F}(t) \equiv \boldsymbol{r_R}(t+52) + c\boldsymbol{s}(t)\boldsymbol{\varepsilon}, \tag{6.4.1}$$

$$\boldsymbol{\Sigma_F}(t) \equiv \boldsymbol{\Sigma_R}(t), \tag{6.4.2}$$

where $\boldsymbol{r_R}(t)$ is the realized 52-week return until weekend t, $\boldsymbol{\Sigma_R}(t)$ is the realized covariance matrix of the sectoral returns for the preceding 52 weeks until weekend t,

$$\boldsymbol{s}(t) \equiv \begin{pmatrix} s_1(t) & & 0 \\ & \ddots & \\ 0 & & s_{10}(t) \end{pmatrix}, \boldsymbol{\varepsilon} \equiv \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_{10} \end{pmatrix}, \tag{6.4.3}$$

$s_i(t)$ is the standard deviations of each sector's return for the preceding 52 weeks until weekend t, $\boldsymbol{z}_i$ are the random variables under $\boldsymbol{N}(0,1)$, each pair $\boldsymbol{z}_i$ and $\boldsymbol{z}_j$ is independent, and c is a scalar coefficient. We realize that c represents a power of the forecast returns in our numerical examples. In this section, predictability-based measurements indicate the predictability-based measurements for the information ratio.

We calculate the test return vectors, the test covariance matrices, the realized return vectors, and the realized covariance matrices for year-end weekends from the first weekend of 2007 to the last weekend of 2016 for $c = 0.1, 0.3,$ or 0.5. In each case, the random variables are given using the random Excel function. Furthermore, in each case, we perform the attribution analysis for the test returns, the forecast magnitude of the test risks, and the test risk factors using Property 6.2. Table 6.1 summarizes the attribution analysis.

Table 6.1: Attribution of Information Ratio to Predictability of Forecast Return and Risk

(A)		(B)		(C)	(D)		(E)											
Year end	Realized Information Ratio			Market Condition	The Limit by the perfect Forecast (%)	The Predictability-Based Measurement of the Forecast												
									(F) Forecast Rerutn			(G) Forecast Risk	(H) Magni -tude Fore- cast	(I) Factor Fore- cast	(X) Dupli- cate Term X	(Y) Duplicate Term Y		
	c =					c =			c =							c =		
	0.1	0.3	0.5			0.1	0.3	0.5	0.1	0.3	0.5					0.1	0.3	0.5
2016	0.192	0.058	0.097	0.562	62.7%	0.308	0.186	0.085	0.374	0.138	0.087	0.434	0.564	0.452	0.046	0.126	0.242	0.191
2015	0.159	0.060	0.083	0.416	94.7%	0.305	0.080	0.207	0.508	-0.005	0.246	0.822	0.816	0.824	0.129	-0.078	0.211	0.086
2014	0.109	0.040	0.010	0.400	97.9%	0.272	0.097	0.109	0.334	0.370	0.164	0.761	0.966	0.801	-0.027	0.156	-0.055	0.163
2013	0.293	0.082	0.118	0.597	80.2%	0.493	0.446	0.222	0.619	0.161	0.045	0.551	0.795	0.531	0.027	0.124	0.536	0.428
2012	0.357	0.212	0.109	0.675	80.0%	0.533	0.169	0.230	0.528	0.135	0.223	0.717	0.840	0.697	-0.020	0.089	0.117	0.091
2011	0.243	0.113	0.105	0.363	95.0%	0.649	0.303	0.373	0.441	0.149	-0.050	0.813	0.830	0.828	0.106	0.344	0.291	0.560
2010	0.124	0.080	-0.020	0.504	84.6%	0.265	0.196	-0.004	0.330	0.119	0.027	0.643	0.980	0.606	-0.097	0.138	0.281	0.173
2009	0.147	0.080	0.073	0.482	90.2%	0.283	0.071	0.076	0.288	0.142	0.022	0.727	0.925	0.754	-0.049	0.171	0.104	0.229
2008	0.198	0.088	0.082	0.469	95.2%	0.345	0.145	0.016	0.353	0.223	0.105	0.600	0.928	0.647	-0.022	0.344	0.274	0.263
2007	0.133	0.083	0.030	0.468	87.1%	0.280	0.053	-0.018	0.458	0.261	0.080	0.595	0.633	0.636	0.196	0.098	0.068	0.179

Data Source: Ten sector total indices. S&P Dow Jones Indices. 2018. S&P Dow Jones Indices LLC.

We consider the first row in Table 6.1. Column (B) indicates the realized information ratio $IR(\boldsymbol{r_F}, \boldsymbol{\Sigma_F} \mid \boldsymbol{r_R}, \boldsymbol{\Sigma_R})$ of (6.1.8), which is the optimal portfolio achieved by the test distribution. We realize that the optimal portfolios based on the test distribution achieve 0.192 of the realized information ratio for the test distribution for the end of 2016 when $c = 0.1$. Column (C) indicates that the Sharpe ratio for the market condition is 0.562, on average. This result indicates that the information ratio does not exceed 0.562 through a non-leveraged long/short portfolio on average. We find that the perfect forecast achieves 62.7 percent of the upper bound of 0.562, on average and that the predictability-based measurement of the test distribution is 0.308 on average in Column (E) when $c = 0.1$. Moreover, the natural outcome is that the optimal portfolio tends to achieve a higher information ratio in Column (B) when the predictability-based measurement of the forecast is higher in Column (E). Column (F) indicates the predictability-based measurement of the test return $P\_IR(\boldsymbol{r_F}, \boldsymbol{\sigma_R}, \boldsymbol{K_R})$ of (6.2.2). Column (G) indicates the predictability-based measurement of the test risk $P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_F})$. Indeed, the predictability-based measurement of the test distribution in Column (E) is attributed to the predictability-based measurement of the test returns and of the test risks without the duplicate term Y in Column (Y) from (6.2.9). We find that the predictability-based measurement of the test return is 0.374 and that the predictability of the test risk is 0.434 when $c = 0.1$. We can simultaneously obtain quantitative information on the predictability-based measurements of the return and risk. From this, we can fairly and directly compare the predictability-based measurements of the return and the risk. The magnitude of the duplicate is 12.6 percent when $c = 0.1$ from Column (Y). Column (H) indicates the predictability-based measurement of the magnitude of the test risk $P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_F}, \boldsymbol{K_R})$ of (6.2.4). Column (I) shows the predictability-based measurement of the test risk factors $P\_IR(\boldsymbol{r_R}, \boldsymbol{\sigma_R}, \boldsymbol{K_F})$ of (6.2.5). The predictability-based measurement of the test risk in Column (G) is attributed to the predictability-based measurement of the magnitude of the test risk and of the test risk factors without the duplicate term X in Column (X) from (6.2.10). We find that the predictability-based measurement of the magnitude of the test risk is 0.564 and that the predictability-based measurement of the test risk factors is 0.452. From this, we can identify the main source of the reduction in the predictability-based measurement of the risk. We realize that the reduction in the predictability-based measurement of the test risk factors relatively causes a reduction in the predictability-based measurement of the test risk in Columns (G), (I), and (I) of the historical data. We note that the situations are different from the case of the realized Sharpe ratio. From these attribution analyses, we can compare and quantitatively identify the main source of the reduction in the realized information ratio. This information may be useful for improving subsequent investments.

Table 6.2: Attribution of Information Ratio to Predictability of Forecast Return and Risk

(In the case that the risk forecast has information on the realized risk for a half period in the future)

(A) Year end	(B) Realized Information Ratio c = 0.1	(B) c = 0.3	(B) c = 0.5	(C) Market Condition	(D) The Limit by the perfect Forecast (%)	(E) The Predictability-Based Measurement of the Forecast c = 0.1	(E) c = 0.3	(E) c = 0.5	(F) Forecast Rerutn c = 0.1	(F) c = 0.3	(F) c = 0.5	(G) Forecast Risk	(H) Magni-tude Fore-cast	(I) Factor Fore-cast	(X) Dupli-cate Term X	(Y) Duplicate Term Y c = 0.1	(Y) c = 0.3	(Y) c = 0.5
2016	0.192	0.065	0.053	0.562	62.7%	0.327	0.204	0.081	0.303	0.117	0.122	0.461	0.447	0.475	0.166	0.191	0.253	0.125
2015	0.206	0.083	0.117	0.416	94.7%	0.379	0.240	0.023	0.265	0.260	0.149	0.768	0.794	0.800	0.122	0.293	0.159	0.053
2014	0.054	0.073	0.016	0.400	97.9%	0.319	0.143	0.132	0.507	0.389	0.270	0.836	0.976	0.874	-0.035	-0.045	-0.104	0.004
2013	0.268	0.146	0.056	0.597	80.2%	0.535	0.316	0.344	0.387	0.162	0.286	0.505	0.856	0.529	-0.078	0.445	0.450	0.354
2012	0.322	0.143	-0.025	0.675	80.0%	0.483	0.159	0.128	0.493	0.305	0.001	0.759	0.856	0.771	-0.069	0.031	-0.105	0.168
2011	0.086	0.069	0.105	0.363	95.0%	0.631	0.224	0.331	0.621	0.051	0.037	0.825	0.849	0.847	0.078	0.135	0.298	0.419
2010	0.104	0.053	0.044	0.504	84.6%	0.256	0.074	0.016	0.403	0.152	0.029	0.693	0.973	0.710	-0.144	0.007	0.075	0.140
2009	0.178	0.119	0.099	0.482	90.2%	0.450	0.058	0.085	0.400	0.282	0.104	0.834	0.844	0.814	0.078	0.119	-0.156	0.050
2008	0.190	0.082	0.055	0.469	95.2%	0.442	0.217	0.286	0.333	0.209	0.064	0.882	0.935	0.919	-0.019	0.179	0.079	0.293
2007	0.188	0.006	-0.030	0.468	87.1%	0.097	0.176	0.341	0.316	0.400	0.021	0.677	0.687	0.676	0.185	-0.026	-0.031	0.514

Data Source: Ten sector total indices. S&P Dow Jones Indices. 2018. S&P Dow Jones Indices LLC.

Table 6.2 indicates the case that the risk forecast has information on the realized risk for half a period in the future. Indeed, we assume that the next 52 weeks' test return vector $\boldsymbol{r_F}(t)$ and the test covariance matrix $\boldsymbol{\Sigma_F}(t)$ for weekend t are calculated as follows:

$$\boldsymbol{r_F}(t) \equiv \boldsymbol{r_R}(t+52) + c\boldsymbol{s}(t)\boldsymbol{\varepsilon}, \tag{6.4.4}$$

$$\boldsymbol{\Sigma_F}(t) \equiv \boldsymbol{\Sigma_R}(t+26), \tag{6.4.5}$$

where $\boldsymbol{r_R}(t)$ is the realized 52-week return until weekend t, $\boldsymbol{\Sigma_R}(t)$ is the realized covariance matrix of the sectoral returns for the preceding 52 weeks until weekend t,

$$\boldsymbol{s}(t) \equiv \begin{pmatrix} s_1(t) & & 0 \\ & \ddots & \\ 0 & & s_{10}(t) \end{pmatrix}, \boldsymbol{\varepsilon} \equiv \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_{10} \end{pmatrix}, \tag{6.4.6}$$

$s_i(t)$ is the standard deviations of each sector's return for the preceding 52 weeks until weekend t, $\boldsymbol{z}_i$ represent the random variables under $\boldsymbol{N}(0,1)$, each $\boldsymbol{z}_i$ and $\boldsymbol{z}_j$ pair is independent, and c is a scalar coefficient. In our numerical examples, we realize that c represents the predictability of the forecast returns.

We find that the predictability-based measurement of the risk slightly enhances a comparison with the case of Table 7.1. Table 7.1 shows the relative enhancement in the predictability of the risk factors.

6.5 Summary

In this chapter, we provide a formula for the ex-post realized information ratio achieved by the optimal portfolio for the information ratio based on the ex-ante forecast return distribution. We propose a measurement that evaluates the power of an ex-ante forecast that predicts the ex-post realized information ratio, such as for the Sharpe ratio. The measurement is called the predictability-based measurement for the information ratio. Additionally, we find that the predictability-based measurement of the perfect forecast is not necessarily one. This result provides insights into the performance analysis of investments. Although the information ratio is an important risk-adjusted performance measure for evaluating investments, this result indicates that applying other performance measures, such as the Sharpe ratio in addition to the information ratio, is desirable in evaluating the predictive power of a forecast when comparing historically different market conditions. Furthermore, we offer an attribution method of the realized information ratio that evaluates the ex-ante forecast risk and return. The attribution formula decomposes the ex-post realized information ratio into the realized market conditions, the predictability-based measurement of the returns, the predictability-based measurement of the magnitude of the risk, and the predictability-based measurement of the risk factors. Numerical examples of the attribution method are provided that the predicting power of the risk magnitude for the information ratio is smaller than the predicting power of the risk magnitude for the Sharpe ratio. We find that perfect forecasts provide information ratios as high as approximately ninety percent of the maximum information ratio in the examples.

Chapter 7

Conclusions and future research

The first contribution of this thesis is to provide formulas that express relations between the forecasts at that moment of portfolio selection and performance measurements. The second contribution is to provide an ex-post realized Sharpe ratio attribution and an ex-post realized information ratio of the ex-ante forecast returns and risk. The third contribution is to provide attributions using formulas that enable us to provide various quantitative performance analyses. The formulas are of the ex-post realized Sharpe ratio achieved by the optimal portfolio based on the ex-ante forecast return distribution and of the ex-post realized information ratio achieved by the optimal portfolio based on the ex-ante forecast return distribution. Furthermore, we offer an attribution method of the realized Sharpe ratio and realized information ratio. The methods compare and quantitatively evaluate the ex-ante forecast returns and risk. Important formulas are illustrated in the following tables.

Table 7.1 Summary of formulas

<table>
<tr><th></th><th></th><th>Sharpe ratio</th><th>Information ratio</th></tr>
<tr><td>1</td><td>Ex-ante forecast return distribution</td><td colspan="2">$\boldsymbol{N}(\boldsymbol{r_F}, \boldsymbol{\Sigma_F})$</td></tr>
<tr><td>2</td><td>Optimal portfolio under ex-ante forecast return distribution</td><td>$\boldsymbol{w_F} = \boldsymbol{\Sigma_F^{-1}} \frac{\boldsymbol{r_F}}{\beta_F}$</td><td>$\boldsymbol{w_F} = c\boldsymbol{\Sigma_F^{-1}}(\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e})$</td></tr>
<tr><td>3</td><td>Ex-post realized return distribution</td><td colspan="2">$\boldsymbol{N}(\boldsymbol{r_R}, \boldsymbol{\Sigma_R})$</td></tr>
<tr><td>4</td><td>Optimal portfolio under ex-post realized return distribution</td><td>$\boldsymbol{w_R} = \boldsymbol{\Sigma_R^{-1}} \frac{\boldsymbol{r_R}}{\beta_R}$</td><td>$\boldsymbol{w_R} = c\boldsymbol{\Sigma_R^{-1}}(\alpha_R \boldsymbol{r_R} - \beta_R \boldsymbol{e})$</td></tr>
<tr><td>5</td><td>Equivalent return[(1)] under ex-post realized risk</td><td>$\boldsymbol{K_R \sigma_R \Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}$</td><td>$\frac{\delta_R}{\alpha_R \delta_F} \boldsymbol{K_R \sigma_R \Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1}}(\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e}) + \frac{\beta_R}{\alpha_R} \boldsymbol{e}$</td></tr>
<tr><td>6</td><td>Maximum Sharpe ratio</td><td colspan="2">$\left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\|$</td></tr>
<tr><td>7</td><td>Ratio achieved by optimal portfolio based on ex-ante forecast return under ex-post realized return</td><td>$\left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\|(\boldsymbol{h}, \boldsymbol{g})$</td><td>$\left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\|(\boldsymbol{f}, \boldsymbol{g})$</td></tr>
<tr><td>8</td><td rowspan="2">Normalized vectors</td><td>$\boldsymbol{h} \equiv \frac{\boldsymbol{\Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}}{\left\|\boldsymbol{\Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}\right\|}$</td><td>$\boldsymbol{f} \equiv \frac{\boldsymbol{\Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1}}(\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e})}{\left\|\boldsymbol{\Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1}}(\alpha_F \boldsymbol{r_F} - \beta_F \boldsymbol{e})\right\|}$</td></tr>
<tr><td>9</td><td colspan="2">$\boldsymbol{g} \equiv \frac{\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}}{\left\|\boldsymbol{\sigma_R^{-1} K_R^{-1} r_R}\right\|}$</td></tr>
</table>

$\alpha_F \equiv \boldsymbol{e}'\boldsymbol{\Sigma_F}^{-1}\boldsymbol{e}, \beta_F \equiv \boldsymbol{e}'\boldsymbol{\Sigma_F}^{-1}\boldsymbol{r_F}, \gamma_F \equiv \boldsymbol{r}'_{\boldsymbol{F}}\boldsymbol{\Sigma_F}^{-1}\boldsymbol{r_F}, \delta_F \equiv \alpha_F\gamma_F - \beta_F^2, \alpha_R \equiv \boldsymbol{e}'\boldsymbol{\Sigma_R}^{-1}\boldsymbol{e}, \beta_R \equiv \boldsymbol{e}'\boldsymbol{\Sigma_R}^{-1}\boldsymbol{r_R}, \gamma_R \equiv \boldsymbol{r_R}'\boldsymbol{\Sigma_R}^{-1}\boldsymbol{r_R}, \delta_R \equiv \alpha_R\gamma_R - \beta_R^2$

$\boldsymbol{\Phi} \equiv \boldsymbol{\sigma_R K_R^{-1} K_F \sigma_F}^{-1}, \boldsymbol{K_F^{-1}\Sigma_F K_F} = \boldsymbol{\sigma_F}^2, \boldsymbol{K_R^{-1}\Sigma_R K_R} = \boldsymbol{\sigma_R}^2$

(1): The equivalent return produces an optimal portfolio identical to the optimal portfolio based on the ex-ante forecast return distribution in the case of the Shape ratio. Additionally, the equivalent return produces an optimal portfolio that achieves an information ratio identical to the information ratio achieved by the optimal portfolio under the ex-ante forecast return distribution for the information ratio.

Row 5: An equivalent return under the ex-post realized risk. Row 7, Row 8, and Row 9: The ratios achieved by the optimal portfolio based on the ex-ante forecast return under the ex-post realized return and representing important formulas introduced in this thesis.

Finally, we decompose the realized Sharpe ratio into the absolute value of the market Sharpe ratio, the predictability-based measurement of the returns, and the predictability-based measurement of the risk, which is decomposed as the predictability-based measurement of the magnitude of the risk and the predictability-based measurement of the risk factors. We also show that the realized information ratio is decomposed into the absolute value of the market Sharpe ratio, the limit by the perfect forecast, the predictability-based measurement of the returns, and the predictability-based measurement of the risk, which is decomposed as the predictability-based measurement of the magnitude of the risk and the predictability-based measurement of the risk factors. Additionally, we illustrate numerical examples of the attribution analysis.

We find that the attribution given by the formula quantitatively enables a comparative analysis between the predictability-based measurement of the returns and the predictability-based measurement of the risk. Furthermore, we find that the predictability-based measurement of the risk has more detailed predictability-based measurements, which are the magnitude of the risk and the risk factors. Additionally, we simultaneously evaluate the quality of the investor's portfolio and the quality of the benchmark in the attribution analysis against the benchmark. In addition, we realize that these attribution analyses provide rich information for an investment analysis. Therefore, these analyses may become useful analytical tools for portfolio managers and sponsors.

However, additional research should be performed on theoretical issues and on applications of the properties discussed in this thesis. For theoretical aspects, the first issue is the relation of ex-ante forecast distribution and ex-post realized distribution. An ex-post realized distribution is given independently with the ex-ante forecast distribution. Indeed, the ex-ante forecast distribution has no information of the ex-post realized distribution. This is the first step of the research. It is natural that the ex-ante distribution and ex-post distribution have some information with each other or that return distributions has a time series structure. However, there are researches that historical returns have little information of future returns. Therefore, risk structure, risk magnitude and risk factors which are important role in this thesis, is an object of the research. A direction of research is an estimation of a stochastic model of the risk structure. Another direction of research is a development of monitoring the behavior of risk structure in short term. Next, analysis of effects of the change of risk structure to the measurement is needed. These researches need empirical studies using historical data. These researches may give methods which prevent from reduction of risk adjusted performance measurement like Sharpe ratio and information ratio. Additionally, they may give ideas of new investment products.

For applied aspects, a discussion and analysis of differences between assumptions in this thesis and facts in investment practices are needed. Normal distributions of returns of investment assets and full rank of covariance matrices of risk are supposed. S&P sector total return indices are constructed from returns of the 500 assets which are well traded and well monitored by market participants. Additionally, noises from individual assets are eliminated from the sector indices by summing up of the return of assets. Therefore, the empirical data using in this thesis tend to satisfy the assumptions. Analysis and evaluation of effects where returns are different from normal distribution are needed. Additionally, enhancement of discussion treating investment assets, which are less liquid or distress, is an issue in application of investment practices. A framework which explains returns of investment assets at common risk factors and others may be needed. These researches may give more insights in investment practices.

Finally, further research is required to develop effective ways to use these properties to analyze different aspects of the investment process and, thereby, enhance the realized Sharpe ratio.

Acknowledgment

The author is grateful to Professor Yumiharu Nakano, Graduate School of Mathematical and Computing Science, Tokyo Institute of Technology, for suggestions and comments that improved the content of this paper.

Appendix 1: The optimal portfolio to maximize Sharpe ratio

We review the derivations in accordance with Shimizu (2017). First, for a given number h, we find the minimum-variance portfolio as the solution to the following quadratic programming problem:

$$minimize\ \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w},$$

$$\text{s.t.}\, r'\boldsymbol{w} = h, \text{and } \boldsymbol{e}'\boldsymbol{w} = 1, \tag{A1.1}$$

where $\boldsymbol{e} = (1, \cdots, 1)'$. Problem (1) has the following unique solution (Steinbach, 2001)

$$\bar{\boldsymbol{w}} = \boldsymbol{\Sigma}^{-1}(\lambda\boldsymbol{e} + k\boldsymbol{r}), \tag{A1.2}$$

where

$$\lambda \equiv \frac{\gamma - \beta h}{\delta}, k \equiv \frac{\alpha h - \beta}{\delta}, \tag{A1.3}$$

$$\alpha \equiv \boldsymbol{e}'\boldsymbol{\Sigma}^{-1}\boldsymbol{e}, \beta \equiv \boldsymbol{e}'\boldsymbol{\Sigma}^{-1}\boldsymbol{r}, \gamma \equiv \boldsymbol{r}\boldsymbol{\Sigma}^{-1}\boldsymbol{r}, \delta \equiv \alpha\gamma - \beta^2. \tag{A1.4}$$

Thus, we obtain the risk of the portfolio given by

$$\sigma_{\bar{\boldsymbol{w}}}^2 = \bar{\boldsymbol{w}}'\boldsymbol{\Sigma}\bar{\boldsymbol{w}} = \frac{\gamma - 2\beta h + \alpha h^2}{\delta}. \tag{A1.5}$$

The Sharpe ratio of the portfolio is given by

$$SR_{\bar{\boldsymbol{w}}} = \frac{h}{\sigma_{\bar{\boldsymbol{w}}}} = \frac{h}{\sqrt{\frac{\gamma - 2\beta h + \alpha h^2}{\delta}}}. \tag{A1.6}$$

We can obtain the maximum $f(h)$ when $h = \gamma/\beta$ holds under the condition that β is positive. By substituting $h = \gamma/\beta$ with (A1.3), we obtain the optimal portfolio given from (A1.2) using

$$\boldsymbol{w}_0 = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{r}}{\boldsymbol{e}'\boldsymbol{\Sigma}^{-1}\boldsymbol{r}}. \tag{A1.7}$$

We obtain a greater $f(h)$ when h becomes larger under the condition that β is not positive. In this case, we continue the research by substituting $\tilde{\boldsymbol{r}}_{\boldsymbol{F}} \equiv \alpha\boldsymbol{r}_{\boldsymbol{F}} - \beta\boldsymbol{e}$ for $\boldsymbol{r}_{\boldsymbol{F}}$. The condition holds for which the sum of the Sharpe ratios of all risk factors becomes negative. However, predicting such a forecast from an economic thought perspective is difficult. The optimal portfolio often has highly leveraged positions and large short positions when h is large. Typically, investors do not hold such extreme portfolios.

Appendix 2: Derivations

A2.1 Derivation of Theorem 3.1

We review the derivations in accordance with Shimizu (2017). From (3.2.2) and (3.2.3), we have

$$\boldsymbol{w}_F = \frac{\boldsymbol{\Sigma_F^{-1} r_F}}{\boldsymbol{e'\Sigma_F^{-1} r_F}} = \frac{\boldsymbol{K_F \sigma_F}^{-2} \boldsymbol{K_F^{-1} r_F}}{\boldsymbol{e' K_F \sigma_F}^{-2} \boldsymbol{K_F^{-1} r_F}}. \tag{A2.1.1}$$

From (3.2.2), (3.2.4), and the definition of $\boldsymbol{\Phi}$, we have

$$\begin{aligned} \boldsymbol{w}_0 &= \frac{\boldsymbol{\Sigma_R^{-1} K_R \sigma_R \Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}}{\boldsymbol{e'\Sigma_R^{-1} K_R \sigma_R \Phi \sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}} \\ &= \frac{\boldsymbol{K_R \sigma_R}^{-2} \boldsymbol{K_R^{-1} K_R \sigma_R \sigma_R K_R^{-1} K_F \sigma_F}^{-1} \boldsymbol{\sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}}{\boldsymbol{e' K_R \sigma_R}^{-2} \boldsymbol{K_R^{-1} K_R \sigma_R \sigma_R K_R^{-1} K_F \sigma_F}^{-1} \boldsymbol{\sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}}. \end{aligned} \tag{A2.1.2}$$

From $\boldsymbol{K_R \sigma_R}^{-2} \boldsymbol{K_R^{-1} K_R \sigma_R \sigma_R K_R^{-1}} = \boldsymbol{I}$, we have

$$\boldsymbol{w}_0 = \frac{\boldsymbol{K_F \sigma_F}^{-1} \boldsymbol{\sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}}{\boldsymbol{e' K_F \sigma_F}^{-1} \boldsymbol{\sigma_F}^{-1} \boldsymbol{K_F^{-1} r_F}} = \boldsymbol{w}_F. \tag{A2.1.3}$$

A2.2 Implied forecast return of a given portfolio

Let $\boldsymbol{w}$ denote the given portfolio, $\boldsymbol{\Sigma_F}$ the forecast covariance matrix, and $\boldsymbol{r}$ the implied forecast return vector,

$$\begin{aligned} &minimize\ \frac{1}{2}\boldsymbol{w'\Sigma_F w} - \mu \boldsymbol{r'_F w}, \\ &\text{s.t.}\ \boldsymbol{e'w} = 1, \end{aligned} \tag{A2.2.1}$$

We solve this problem using the Lagrange multiplier method as follows:

$$\mathrm{L}(\boldsymbol{w}, l) = \frac{1}{2}\boldsymbol{w'\Sigma_F w} - \mu \boldsymbol{r'_F w} - l(\boldsymbol{e'w} - 1) \tag{A2.2.2}$$

Solving

$$\frac{\partial}{\partial \boldsymbol{w}}\left(\frac{1}{2}\boldsymbol{w'\Sigma_F w} - \mu \boldsymbol{r'_F w} - l(\boldsymbol{e'w} - 1)\right) = \boldsymbol{0}, \tag{A2.2.3}$$

$$\frac{d}{dl}\left(\frac{1}{2}\boldsymbol{w'\Sigma_F w} - \mu \boldsymbol{r'_F w} - l(\boldsymbol{e'w} - 1)\right) = 0, \tag{A2.2.4}$$

Then

$$\frac{\partial}{\partial \boldsymbol{w}}\left(\frac{1}{2}\boldsymbol{w'\Sigma_F w} - \mu \boldsymbol{r'_F w}\right) = \boldsymbol{0}. \tag{A2.2.5}$$

We have

$$\boldsymbol{r_F} = \frac{1}{\mu}\boldsymbol{\Sigma_F w}. \tag{A2.2.6}$$

We let $\mu \equiv 1$ because $\boldsymbol{r_F}$ is subsequently normalized.

A2.3 Minimum variance portfolio

As with Appendix 1, by solving the quadratic programming problem (A1.1), we obtain the risk of the portfolio given by

$$\sigma_{\overline{\boldsymbol{w}}}^2 = \overline{\boldsymbol{w}}'\boldsymbol{\Sigma_F}\overline{\boldsymbol{w}} = \frac{\gamma - 2\beta h + \alpha h^2}{\delta}. \tag{A2.3.1}$$

By solving $\frac{d}{dh}\sigma_{\overline{\boldsymbol{w}}}^2 = \frac{2}{\delta}(-\beta + \alpha h) = 0$, we obtain $h = \beta/\alpha$. From (A1.2) and (A1.3), we obtain the weight vector $\boldsymbol{w}_{min}$ of the minimum variance portfolio

$$\boldsymbol{w}_{min} = \frac{1}{\alpha}\boldsymbol{\Sigma}_F^{-1}\boldsymbol{e}. \tag{A2.3.2}$$

A2.4 Derivations for the fundamental law of active management

In this appendix, we review the fundamental law of active management in accordance with Grinold and Kahn (1999).

Suppose that the residual returns for n risky assets $\boldsymbol{\theta}$

$$\boldsymbol{\theta} = \boldsymbol{Ax} \tag{A2.4.1}$$

where

$$\boldsymbol{x} \equiv \begin{pmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix}, \tag{A2.4.2}$$

$\boldsymbol{x}_i$ is the random variables under a normal distribution $\boldsymbol{N}(0,1)$, and each pair $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ is independent.

Suppose that an investor has BR signals $\boldsymbol{z} \equiv \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_{\mathrm{BR}} \end{pmatrix}$ for the residual return $\boldsymbol{\theta}$ and that

$$\boldsymbol{z} = \boldsymbol{Ey} \tag{A2.4.3}$$

where

$$\boldsymbol{y} \equiv \begin{pmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{\mathrm{BR}} \end{pmatrix}, \tag{A2.4.4}$$

$\boldsymbol{y}_i$ represents the BR random variables under $\boldsymbol{N}(0,1)$, and each pair $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ is independent. Let $\boldsymbol{E}$ = the square root of the covariance matrix of $\boldsymbol{z}$ and let $\boldsymbol{J}$ be the inverse of $\boldsymbol{E}$.

$$\mathrm{Var}(\boldsymbol{z}) = \boldsymbol{E}\boldsymbol{E}' \tag{A2.4.5}$$

where $\boldsymbol{E}'$ denotes the transposed matrix of $\boldsymbol{E}$.

Let

$$\boldsymbol{Q} = \mathrm{Cov}(\boldsymbol{\theta}, \boldsymbol{z}) \tag{A2.4.6}$$

be the covariance matrix, which has N rows and BR colmns.

Let

$$\boldsymbol{P} = \mathrm{Cov}(\boldsymbol{x}, \boldsymbol{y}) \tag{A2.4.7}$$

be the covariance matrix. It follows that

$$\boldsymbol{Q} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{E}'. \tag{A2.4.8}$$

The mean and variance of $\boldsymbol{\theta}$ conditional on the signal $\mathrm{Cov}(\boldsymbol{\theta}, \boldsymbol{z})$

$$\mathrm{E}(\boldsymbol{\theta} | \, \boldsymbol{z}) = \boldsymbol{\alpha}(\boldsymbol{z}) = \boldsymbol{A}\boldsymbol{P}\boldsymbol{E}^{-1}\boldsymbol{z} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{J}\boldsymbol{z} \tag{A2.4.9}$$

Furthermore, we have

$$\mathrm{Var}(\boldsymbol{\theta} | \, \boldsymbol{z}) = \mathbf{G} = \mathbf{A}(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}')\boldsymbol{A}' \tag{A2.4.10}$$

For a given signal $\boldsymbol{z}$, we find the optimal portfolio $\mathrm{h}^*(\boldsymbol{z})$ as the solution to the following quadratic programming problem:

$$maximize\ \mathrm{h}(\boldsymbol{z})\boldsymbol{\alpha}(\boldsymbol{z}) - \lambda\, \mathrm{h}(\boldsymbol{z})'\mathbf{G}\mathrm{h}(\boldsymbol{z}) \tag{A2.4.11}$$

The first-order conditions for the maximization problem (A2.4.11) are

$$\boldsymbol{\alpha}(\boldsymbol{z}) = 2\lambda\, \mathbf{G}\mathrm{h}(\boldsymbol{z}) \tag{A2.4.12}$$

Substituting (A2.4.9) and (A2.4.10) into (A2.4.12) gives

$$\boldsymbol{A}\boldsymbol{P}\boldsymbol{J}\boldsymbol{z} = 2\lambda\mathbf{A}(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}')\boldsymbol{A}'\mathrm{h}^*(\boldsymbol{z}) \tag{A2.4.13}$$

Therefore, we have

$$\boldsymbol{A}'\mathrm{h}^*(\boldsymbol{z}) = \frac{1}{2\lambda}(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}')^{-\mathbf{1}}\boldsymbol{P}\boldsymbol{J}\boldsymbol{z} \tag{A2.4.14}$$

From $\boldsymbol{P}$, which is a covariance matrix, we have

$$\mathrm{h}^*(\boldsymbol{z})'\boldsymbol{A} = \frac{1}{2\lambda}\boldsymbol{z}'\boldsymbol{J}'\boldsymbol{P}'((\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}')^{-\mathbf{1}}). \tag{A2.4.15}$$

Finally, we have

$$\mathrm{h}^*(\boldsymbol{z})' = \frac{1}{2\lambda}\boldsymbol{z}'\boldsymbol{J}'\boldsymbol{P}'((\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}')^{-\mathbf{1}})\boldsymbol{A}^{-\mathbf{1}}. \tag{A2.4.16}$$

From (2.2.9) and (2.2.15), the optimal portfolio's residual return is

$$\mathrm{h}^*(\boldsymbol{z})'\boldsymbol{\alpha}(\boldsymbol{z}) = \frac{1}{2\lambda}\boldsymbol{z'J'P'}((\boldsymbol{I}-\boldsymbol{PP'})^{-\mathbf{1}})\boldsymbol{PJz}. \tag{A2.4.17}$$

From (A2.4.10) and (A2.4.15), the optimal portfolio's residual risk is

$$\mathrm{h}^*(\boldsymbol{z})'\mathrm{G}\mathrm{h}^*(\boldsymbol{z}) = \left(\frac{1}{2\lambda}\right)^2 \boldsymbol{z'J'P'}((\boldsymbol{I}-\boldsymbol{PP'})^{-\mathbf{1}})(\boldsymbol{I}-\boldsymbol{PP'})\boldsymbol{A'}(\boldsymbol{A'})^{-\mathbf{1}}(\boldsymbol{I}-\boldsymbol{PP'})^{-\mathbf{1}}\boldsymbol{PJz}$$

$$= \left(\frac{1}{2\lambda}\right)^2 \boldsymbol{z'J'P'}((\boldsymbol{I}-\boldsymbol{PP'})^{-\mathbf{1}})\boldsymbol{PJz}. \tag{A2.4.18}$$

Therefore, from (A2.4.17) and (A2.4.18), the squared information ratio conditional on signal $\boldsymbol{z}$ will be

$$\mathrm{IR}^2(\boldsymbol{z}) = \boldsymbol{z'J'P'}((\boldsymbol{I}-\boldsymbol{PP'})^{-\mathbf{1}})\boldsymbol{PJz} \tag{A2.4.19}$$

Substituting (2.2.3) into (2.2.19) gives

$$\mathrm{IR}^2(\boldsymbol{z}) = \boldsymbol{y}'(\boldsymbol{P}'((\boldsymbol{I}-\boldsymbol{PP'})^{-\mathbf{1}})\boldsymbol{P})\boldsymbol{y} \tag{A2.4.20}$$

Because $\boldsymbol{y}_i$ represents the random variables under $\boldsymbol{N}(0,1)$ and each pair $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ is independent, the unconditional squared information ratio is

$$\mathrm{IR}^2 = \mathrm{E}\{\mathrm{IR}^2(\boldsymbol{z})\} = \mathrm{Tr}\{\boldsymbol{P}'((\boldsymbol{I}-\boldsymbol{PP'})^{-\mathbf{1}})\boldsymbol{P}\}, \tag{A2.4.21}$$

where $\mathrm{Tr}\{\circ\}$ is the trace (sum of the diagonal elements).

Using matrix inversion approximation

$$(\boldsymbol{I}-\boldsymbol{PP'})^{-\mathbf{1}} = \boldsymbol{I} + (\boldsymbol{PP'}) + (\boldsymbol{PP'PP'}) + \cdots, \tag{A2.4.22}$$

we have

$$\boldsymbol{P}'((\boldsymbol{I}-\boldsymbol{PP'})^{-\mathbf{1}})\boldsymbol{P} = \boldsymbol{PP'} + (\boldsymbol{PP'PP'}) + \cdots \tag{A2.4.23}$$

With typical correlations being extremely close to zero and the most optimistic being close to 0.1, we can safely ignore all but the first term in (2.2.23). Therefore, the trace of PP' is

$$\mathrm{Tr}\{\boldsymbol{P}'((\boldsymbol{I}-\boldsymbol{PP'})^{-\mathbf{1}})\boldsymbol{P}\} \cong \mathrm{Tr}\{\boldsymbol{PP'}\} = \sum_{n=1}^{N}\sum_{b=1}^{BR}\rho_{n,b}^2, \tag{A2.4.24}$$

which sums the correlations between orthonormal basis elements $\boldsymbol{x}$ of the residual returns and independent signals $\boldsymbol{y}$ over all assets and signals.

We define the sum of the correlation of each signal with the basis element over the basis element:

$$\zeta_b^2 \equiv \sum_{n=1}^{N}\rho_{n,b}^2. \tag{A2.4.25}$$

From (A2.4.21) and (A2.4.25), we have

$$\mathrm{IR}^2 = \sum_{b=1}^{BR}\zeta_b^2. \tag{A2.4.26}$$

Finally, by assuming that all of the signals have equal value,

$$\zeta_b^2 = \rho^2 = \mathrm{IC}^2 \qquad \text{(A2.4.27)}$$

we find the results:

$$\mathrm{IR}^2 = \mathrm{BR} \cdot \mathrm{IC}^2 \qquad \text{(A2.4.28)}$$

A2.5 Example of Black-Litterman model

Let $\boldsymbol{P} = (1, 0, \cdots, 0), \boldsymbol{v} = v, \boldsymbol{\Omega} = \omega^2$, and $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_n)'$. Then, we obtain

$$\boldsymbol{\mu_{BL}}(\boldsymbol{v}) = \boldsymbol{\pi} + \boldsymbol{\sigma_R P'}(\boldsymbol{P\sigma_R P'} + \boldsymbol{\Omega})^{-1}(\boldsymbol{v} - \boldsymbol{P\pi}) = \begin{pmatrix} \dfrac{\omega^2\pi_1 + \sigma_R^2(1)v}{\sigma_R^2(1) + \omega^2} \\ \pi_2 \\ \vdots \\ \pi_n \end{pmatrix}, \qquad \text{(A2.5.1)}$$

$$\boldsymbol{\Sigma_{BL}} = \boldsymbol{\sigma_R} - \boldsymbol{\sigma_R P'}(\boldsymbol{P\sigma_R P'} + \boldsymbol{\Omega})^{-1}\boldsymbol{P\sigma_R}$$

$$= \begin{pmatrix} \dfrac{\sigma_R^2(1)\omega^2}{\sigma_R^2(1) + \omega^2} & 0 & \cdots & 0 \\ 0 & \sigma_R^2(2) & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & \sigma_R^2(n) \end{pmatrix}. \qquad \text{(A2.5.2)}$$

A2.6 Maximizing the information ratio

We construct the optimal active portfolio $\boldsymbol{w_A}$ that achieves the maximum information ratio based on the return distribution $\boldsymbol{N}(\boldsymbol{r}, \boldsymbol{\Sigma})$ by solving the following quadratic programming problem.

Problem

$$\begin{gathered} minimize\ \boldsymbol{w_A}'\boldsymbol{\Sigma w_A} \\ \text{s.t.}\ \boldsymbol{r}'\boldsymbol{w_A} = h \text{ and } \boldsymbol{e}'\boldsymbol{w_A} = 0\,. \end{gathered} \qquad \text{(A2.6.1)}$$

We solve this problem using the Lagrange multiplier method as follows:

$$L(\boldsymbol{w}, \lambda, k;\ h) = \boldsymbol{w_A}'\boldsymbol{\Sigma w_A} - \lambda\boldsymbol{e}'\boldsymbol{w_A} - k(\ \boldsymbol{r}'\boldsymbol{w_A} - h). \qquad \text{(A2.6.2)}$$

Solving

$$\frac{\partial}{\partial \boldsymbol{w}}\left(\boldsymbol{w_A}'\boldsymbol{\Sigma}\boldsymbol{w_A} - \lambda\boldsymbol{e}'\boldsymbol{w_A} - k(\boldsymbol{r}'\boldsymbol{w_A} - h)\right) = 0,$$

$$\frac{\partial}{\partial \lambda}\left(\boldsymbol{w_A}'\boldsymbol{\Sigma}\boldsymbol{w_A} - \lambda\boldsymbol{e}'\boldsymbol{w_A} - k(\boldsymbol{r}'\boldsymbol{w_A} - h)\right) = 0,$$

$$\frac{\partial}{\partial k}\left(\boldsymbol{w_A}'\boldsymbol{\Sigma}\boldsymbol{w_A} - \lambda\boldsymbol{e}'\boldsymbol{w_A} - k(\boldsymbol{r}'\boldsymbol{w_A} - h)\right) = 0. \qquad \text{(A2.6.3)}$$

We obtain the solution:

$$\boldsymbol{w_A} = \boldsymbol{\Sigma}^{-1}(\lambda\boldsymbol{e} + k\boldsymbol{r}), \qquad \text{(A2.6.4)}$$

$$\lambda = \frac{-\beta h}{\delta}, k = \frac{\alpha h}{\delta}, \qquad \text{(A2.6.5)}$$

where

$$\alpha \equiv \boldsymbol{e}'\boldsymbol{\Sigma}^{-1}\boldsymbol{e}, \beta \equiv \boldsymbol{e}'\boldsymbol{\Sigma}^{-1}\boldsymbol{r}, \gamma \equiv \boldsymbol{r}'\boldsymbol{\Sigma}^{-1}\boldsymbol{r}, \delta \equiv \alpha\gamma - \beta^2. \qquad \text{(A2.6.6)}$$

We obtain the risk of the active portfolio against the benchmark given by

$$\sigma^2_{\boldsymbol{w_A}} = \boldsymbol{w}'_{\boldsymbol{A}}\boldsymbol{\Sigma}\boldsymbol{w_A} = \left(\boldsymbol{\Sigma}^{-1}(\lambda\boldsymbol{e} + k\boldsymbol{r})\right)'\boldsymbol{\Sigma}\boldsymbol{w_A}\left(\boldsymbol{\Sigma}^{-1}(\lambda\boldsymbol{e} + k\boldsymbol{r})\right) = \frac{h^2}{\delta}. \qquad \text{(A2.6.7)}$$

Therefore, the optimal active portfolio achieves the maximum information ratio

$$\mathrm{IR}_{\boldsymbol{w_A}} = \frac{h}{\sqrt{\frac{h^2\alpha}{\delta}}} = \sqrt{\frac{\delta}{\alpha}} = \sqrt{\frac{\alpha\gamma - \beta^2}{\alpha}} = \sqrt{\gamma - \frac{\beta^2}{\alpha}}. \qquad \text{(A2.6.8)}$$

Let $c \equiv h/\delta$, and we have the optimal active portfolio

$$\boldsymbol{w_A}(c) = c\boldsymbol{\Sigma}^{-1}(\alpha\boldsymbol{r} - \beta\boldsymbol{e}). \qquad \text{(A2.6.9)}$$

Appendix 3: A property of the equal weight portfolio

A3.1 Number of investment strategies

Let the number of assets be N and let T be the number of times that the portfolio is rebalanced in the time horizon, including at the start. Suppose that the weights of the assets in the portfolio at the start and after each rebalancing are given as fractions that have the integer M as a common denominator. Suppose that the portfolio does not have any short or leveraged positions and that we can trade without frictions. N and T are also integers and $N \geq 2$. In addition, suppose that $M \geq N$. Let $t = 0$ at the start. Henceforth, we refer to these conditions by $Cond$.
Given these conditions, we can develop an arbitrary investment strategy at the start and after each rebalancing using a set of weights of the assets

$$\left(\frac{m_1}{M}, \frac{m_2}{M}, \cdots, \frac{m_N}{M}\right), \tag{A3.1}$$

where $\frac{m_i}{M}$ is the weight of the i-th asset in the portfolio, m_i is an integer, $0 \leq m_i \leq M$ for all i, and $m_1 + m_2 + \cdots + m_N = M$. Therefore, the number of investment strategies at the start and after each rebalancing equals the number of possible N-sets as previously described. Using the conditions for m_i as a basis, this number represents the number of possible ways to allocate M "objects" to N "boxes," that is, ${}_{M+N-1}C_{N-1}$.[1]
Let L denote this number. Consequently, we have $L^T = ({}_{M+N-1}C_{N-1})^T$ investment strategies in the time horizon because the number of times that the portfolio is rebalanced in the time horizon is T, including at the start. L^T represents the number of strategies in the time horizon when M, N, and T are given.

A3.2 Equal weight rebalancing as the average of all investment strategies

Each time the portfolio is rebalanced, L strategies exist. When ordering these strategies, let $p_l(t)$ be the performance of the l-th strategy during the rebalancing time t to $t+1$ for $l = 1,2,\cdots,L$. We use the term "performance" to denote the total return. Suppose that we select the $l(t)$-th strategy at rebalancing time t for $t = 0,1,\cdots,T-1$. Then, the performance during the time horizon is as follows:

$$\prod_{t=0}^{T-1} p_{l(t)}(t). \tag{A3.2}$$

We let $\overline{P}$ denote the average performance of all strategies during the time horizon. We have a total of L^T strategies and obtain

$$\overline{P} = \frac{1}{L^T} \sum_{L^T} \prod_{t=0}^{T-1} p_{l(t)}(t), \tag{A3.3}$$

where the summation is for all strategies. Indeed, all accumulated performances during the time horizon for all combinations of strategies, $p_1(0)p_1(1) \cdots p_1(T-1)$, $p_1(0)p_1(1) \cdots p_2(T-1), \dots,$ $p_L(0)p_L(1) \cdots p_{L-1}(T-1)$, and $p_L(0)p_L(1) \cdots p_L(T-1)$, appear in the summation. Therefore, we can switch the positions of the summation and the multiplication and obtain

$$\overline{P} = \prod_{t=0}^{T-1} (\frac{1}{L} \sum_{l=1}^{L} p_l(t)). \tag{A3.4}$$

Let the weight of the i-th asset in the portfolio of the l-th strategy be $\frac{m_{i,l}}{M}$, and let the return of the i-th asset during the rebalancing time t to $t+1$ be $r_i(t)$. We obtain the performance of the l-th strategy during the rebalancing time t to $t+1$ as follows:

$$p_l(t) = \frac{m_{1,l}}{M}\big(1 + r_1(t)\big) + \cdots + \frac{m_{N,l}}{M}\big(1 + r_N(t)\big). \qquad \text{(A3.5))}$$

Substituting this equation into Equation (A3.4), we get

$$\overline{P} = \prod_{t=0}^{T-1} \left(\frac{1}{L} \left(\frac{\sum_{l=1}^{L} m_{1,l}}{M} \big(1 + r_1(t)\big) + \cdots + \frac{\sum_{l=1}^{L} m_{N,l}}{M} \big(1 + r_N(t)\big) \right) \right). \tag{A3.6}$$

We calculate the summations in this equation over m instead of l to get

$$\sum_{l=1}^{L} m_{1,l} = \cdots = \sum_{l=1}^{L} m_{N,l} = \frac{LM}{N}. \tag{A3.7}$$

Therefore, we obtain

$$\overline{P} = \prod_{t=0}^{T-1} \left(\frac{1}{N} \big(1 + r_1(t)\big) + \cdots + \frac{1}{N} \big(1 + r_N(t)\big) \right). \tag{A3.8}$$

The right-hand side of this equation expresses the performance of the equal weight rebalancing strategy. Consequently, we obtain the following property.

Property A3:

Assume that $Cond$ hold. Then, the average performance of all investment strategies equals the performance of the equal weight rebalancing strategy.

Appendix 4: Convergence of Sharpe ratio for multi-period investments

From (3.5.6), the expectation value of the observed Sharpe ratio is the product of the ex-post Sharpe ratio and the coefficient:

$$c(T) \equiv \frac{\left(\frac{T}{2}\right)^{1/2} \boldsymbol{\Gamma}\left(\frac{T-1}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{T}{2}\right)}, \tag{A4.1}$$

Suppose $T = 2q$, where q is an integer.

$$c(T) = c(2q) = \frac{\left(\frac{2q}{2}\right)^{1/2} \boldsymbol{\Gamma}\left(\frac{2q-1}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{2q}{2}\right)}$$

$$= \frac{(q)^{1/2} \boldsymbol{\Gamma}\left(\frac{2q-1}{2}\right)}{\boldsymbol{\Gamma}(q)}$$

$$= \frac{\sqrt{q}\frac{2q-3}{2}\frac{2q-5}{2}\cdots\frac{3}{2}\frac{1}{2}\boldsymbol{\Gamma}\left(\frac{1}{2}\right)}{(q-1)\cdot\cdots\cdot 2\cdot 1\cdot \boldsymbol{\Gamma}(1)}$$

$$= \sqrt{q}\sqrt{\pi}\frac{\frac{2q-3}{2}\frac{2q-5}{2}\cdots\frac{3}{2}\frac{1}{2}}{(q-1)\cdot\cdots\cdot 2\cdot 1}. \tag{A4.2}$$

We have

$$\begin{aligned} &2^{q-2}(q-2)! \\ &= 2^{q-2}(q-2)(q-3)\cdots(2)\cdot(1) \\ &= 2(q-2)2(q-3)\cdots 2\cdot(2)\cdot 2\cdot(1) \\ &= (2q-4)(2q-6)\cdot\cdots\cdot 4\cdot 2. \end{aligned} \tag{A4.3}$$

Therefore,

$$\begin{aligned} &2^{q-2}(q-2)!\{(2q-3)(2q-5)\cdots 3\cdot 1\} \\ &= (2q-3)(2q-4)(2q-5)(2q-6)\cdots 4\cdot 3\cdot 2\cdot 1 \\ &= (2q-3)!. \end{aligned} \tag{A4.4}$$

From (A4.2) and (A4.4),

$$\begin{aligned} &2^{q-1}2^{q-2}(q-2)!\ c(2q) \\ &= \sqrt{q}\sqrt{\pi}\frac{(2q-3)(2q-4)\cdots 3\cdot 2\cdot 1}{(q-1)\cdot\cdots\cdot 2\cdot 1} \\ &= \sqrt{q}\sqrt{\pi}\frac{(2q-3)!}{(q-1)!}. \end{aligned} \tag{A4.5}$$

From (A4.5),

$$c(T) = c(2q) = \frac{1}{2^{q-1}2^{q-2}(q-2)!}\sqrt{q}\sqrt{\pi}\frac{(2q-3)!}{(q-1)!}$$

$$= \sqrt{q}\sqrt{\pi}\frac{1}{2^{q-1}2^{q-2}}\frac{(2q-3)!}{(q-1)!\,(q-2)!}. \tag{A4.6}$$

From Stirling' approximation formula, when $n \to \infty$,

$$n! \to \sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}. \tag{A4.7}$$

Substitute (A4.7) into (A4.6), when $q \to \infty$,

$$c(2q) \to \sqrt{q}\sqrt{\pi}\frac{1}{2^{q-1}2^{q-2}}\frac{\sqrt{2\pi(2q-3)}\left(\frac{2q-3}{e}\right)^{2q-3}}{\sqrt{2\pi(q-1)}\left(\frac{q-1}{e}\right)^{q-1}\sqrt{2\pi(q-2)}\left(\frac{q-2}{e}\right)^{q-2}}. \tag{A4.8}$$

Consequently, when $q \to \infty$,

$$c(2q) \to \sqrt{q}\frac{1}{2^{q-1}2^{q-2}}\frac{\sqrt{(2q-3)}\left(\frac{2q-3}{e}\right)^{2q-3}}{\sqrt{2(q-1)}\left(\frac{q-1}{e}\right)^{q-1}\sqrt{(q-2)}\left(\frac{q-2}{e}\right)^{q-2}}$$

$$= \sqrt{\frac{q(2q-3)}{2(q-1)(q-2)}}\frac{1}{2^{q-1}2^{q-2}}\frac{(2q-3)^{2q-3}}{(q-1)^{q-1}(q-2)^{q-2}}\frac{\left(\frac{1}{e}\right)^{2q-3}}{\left(\frac{1}{e}\right)^{q-1}\left(\frac{1}{e}\right)^{q-2}}$$

$$= \sqrt{\frac{q(2q-3)}{2(q-1)(q-2)}}\frac{1}{2^{q-1}2^{q-2}}\frac{(2q-3)^{2q-3}}{(q-1)^{q-1}(q-2)^{q-2}}$$

$$= \sqrt{\frac{\left(2-\frac{3}{q}\right)}{2\left(1-\frac{1}{q}\right)\left(1-\frac{2}{q}\right)}}\frac{1}{2^{q-1}2^{q-2}}\frac{(2q-3)^{q-1}(2q-3)^{q-2}}{(q-1)^{q-1}(q-2)^{q-2}}$$

$$= \sqrt{\frac{\left(2-\frac{3}{q}\right)}{2\left(1-\frac{1}{q}\right)\left(1-\frac{2}{q}\right)}}\frac{\left(\frac{2q-3}{2}\right)^{q-1}\left(\frac{2q-3}{2}\right)^{q-2}}{(q-1)^{q-1}(q-2)^{q-2}}$$

$$= \sqrt{\frac{\left(2-\frac{3}{q}\right)}{2\left(1-\frac{1}{q}\right)\left(1-\frac{2}{q}\right)}} \left(\frac{q-\frac{3}{2}}{q-1}\right)^{q-1} \left(\frac{q-\frac{3}{2}}{q-2}\right)^{q-2}$$

$$= \sqrt{\frac{\left(2-\frac{3}{q}\right)}{2\left(1-\frac{1}{q}\right)\left(1-\frac{2}{q}\right)}} \left(1-\frac{\frac{1}{2}}{q-1}\right)^{q-1} \left(1+\frac{\frac{1}{2}}{q-2}\right)^{q-2}$$

$$= \sqrt{\frac{\left(2-\frac{3}{q}\right)}{2\left(1-\frac{1}{q}\right)\left(1-\frac{2}{q}\right)}} \left(1-\frac{1}{\frac{q-1}{2}}\right)^{q-1} \left(1+\frac{1}{\frac{q-2}{2}}\right)^{q-2}$$

$$= \sqrt{\frac{\left(2-\frac{3}{q}\right)}{2\left(1-\frac{1}{q}\right)\left(1-\frac{2}{q}\right)}} \left(\left(1+\frac{1}{-\frac{q-1}{2}}\right)^{-\frac{q-1}{2}}\right)^{-2} \left(\left(1+\frac{1}{\frac{q-2}{2}}\right)^{\frac{q-2}{2}}\right)^{2}. \quad \text{(A4.9)}$$

When $q \to \infty$, the second and the third terms converge e. Consequently, when $q \to \infty$

$$c(T) = c(2q) \to \sqrt{\frac{2}{2}} e^{-2} e^{2} = 1. \quad \text{(A4.10)}$$

Next, suppose $T = 2q - 1$, where $q > 1$ is an integer. From (A4.1)

$$c(T) = c(2q-1) = \frac{\left(\frac{2q-1}{2}\right)^{1/2} \boldsymbol{\Gamma}\left(\frac{2q-1-1}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{2q-1}{2}\right)}$$

$$= \frac{\left(\frac{2q-1}{2}\right)^{1/2} \boldsymbol{\Gamma}(q-1)}{\boldsymbol{\Gamma}\left(\frac{2q-1}{2}\right)}$$

$$= \left(\frac{2q-1}{2}\right)^{1/2} \frac{(q-2)!}{\left(\frac{2q-3}{2}\right)\left(\frac{2q-5}{2}\right)\cdots\left(\frac{1}{2}\right)\boldsymbol{\Gamma}\left(\frac{1}{2}\right)}. \quad \text{(A4.11)}$$

We have

$$(2q-4)(2q-6)\cdots(2) = 2^{q-2}(q-2). \quad \text{(A4.12)}$$

From (14.11) and (A4.12),

$$c(T) = c(2q-1) = \left(\frac{2q-1}{2}\right)^{1/2} \frac{2^{q-1}2^{q-2}(q-2)!\,(q-2)!}{2^{q-1}2^{q-2}(q-2)!\left(\frac{2q-3}{2}\right)\left(\frac{2q-5}{2}\right)\cdots\left(\frac{1}{2}\right)\boldsymbol{\Gamma}\left(\frac{1}{2}\right)}$$

$$= \left(\frac{2q-1}{2}\right)^{1/2} \frac{2^{q-1}2^{q-2}(q-2)!\,(q-2)!}{(2q-3)!\,\boldsymbol{\Gamma}\left(\frac{1}{2}\right)}$$

$$= \left(\frac{2q-1}{2}\right)^{1/2} \frac{2^{q-1}2^{q-2}(q-2)!\,(q-2)!}{\sqrt{\pi}(2q-3)!}. \tag{A4.13}$$

For examples of q's, the coefficients are following.

Table A4.1 Sample calculation of the coefficient

q	T=2q	coefficient	T=2q-1	coefficient
2	4	1.2533	3	1.3820
3	6	1.1512	5	1.1894
4	8	1.1078	7	1.1259
5	10	1.0837	9	1.0942
6	12	1.0684	11	1.0753
12	24	1.0327	23	1.0342
18	36	1.0215	35	1.0221
30	60	1.0127	59	1.0129

References

Bertsimas, D., et al. (2012). "Inverse Optimization: A New Perspective on the Black-Litterman Model," *Operations Research* 60(6), 1389-1403.

Chan, R. H., et al. (2017). "An Upper Bound for Ex-Post Sharpe Ratio with Application in Performance Measurement," The Journal of Performance measurement 22(1), 7-21.

Chincarini, B. L. and Kim, D. (2013). "Uses and Misuses of the Black-Litterman Model in Portfolio Construction," *Journal of Mathematical Finance* 3, 153-164.

Cover, M. T. (2006). Elements of Information Theory, Second Edition. John Wiley & Sons. Inc., New Jersey.

Elton, J. E. and Gruber, J. M. (1987). Modern Portfolio Theory and Investment Analysis, Third Edition. John Wiley & Sons. Inc., New York.

Farrell, J. L. (1983). Guide to Portfolio Management. McGraw-Hill, Inc.. New York.

Fabozzi, F. J., et al. (2006). "Incorporating Trading Strategies in the Black-Litterman Framework," *The Journal of Trading*, 1(2), 28-37.

Fama, E. and K. French. (1993). "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics,* 33, 3-56.

Grinold, R. C., and Kahn, R. N. (1999). Active Portfolio Management. 2nd ed. McGraw-Hill, New York.

Kelly, J. L. Jr. (1956). "A New Interpretation of Information Rate,) *Bell System Tech. Journal* 35, 917-926.

Lo, A. W. (2002) "The Statistics of Sharpe Ratios," *Financial Analysts Journal*. 59(5), 36-52.

Markowitz, M. H. (1952). "Portfolio Selection," *Journal of Finance* 7(1), 77-91.

Menchero, J. (2016). "Factor Misalignment and Portfolio Construction," *Journal Of Investment Management*. 14(2), 71-85.

Meucci, A. (2010). "The Black-Litterman Approach: Original Model and Extensions," *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1117574 (accessed 15 May 2018).

Meucci, A., 2007. Risk and Asset Allocation. Springer-Verlag, Berlin.

Minotani, C. (2009). Mathematical Statistics Handbook (in Japanese). Mimizuku, Tokyo.

Mittelhammer, R. C. (2013). Mathematical Statistics for Economics and Business, 2nd ed. Springer, New York.

MSCI. (2018). BARRA'S RISK MODEL. https://www.msci.com/www/research-paper/barra-s-risk-models/014972229 (accessed 3rd December 2018)

Owen, D. B. (1968). “A survey of Properties and Applications of the Noncentral *t*-Distribution,” *Technometrics* 10(3), 445-478.

Qian, E. and Gorman S. (2001). “Conditional Distribution in Portfolio Theory,” *Financial Analysts Journal* March/April, 44-51.

R Core Team. (2016). R, A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ (accessed 29 May 2018).

Ross, S.A.. (1976). “The Arbitrage Theory of Capital Asset Pricing, Journal of Economic Theory,” *Journal of Economic Theory* 13(3), 341-360.

Schoettle, K., et al. (2010). “Comparison and Robustification of Bayes and Black-Litterman Models,” *Mathematical Methods of Operations Research* 71, 453-475.

Sharpe, W. F., and Alexander G., 1999. Investments. 4th ed. Prentice Hall. Inc., New Jersey.

Shimizu, M. (Forthcoming). “Attribution of Ex-Post Realized Sharpe Ratio to the Predictability of the Ex-Ante Forecast Return and Risk” *Journal Of Investment Management.*

Shimizu, M. (2017). “A Measure for Evaluating the Distributions of Ex-Ante Forecast Returns,” *Journal of Performance Measurement* 21(4), 52-60.

Shimizu, M. (2010). ”Equally Weighted Rebalancing as the Average of All Investment Strategies”. *Journal Of Investment Management.* 8(4), 27-40.

S&P Dow Jones Indices. (2018). S&P Dow Jones Indices LLC. New York. http://us.spindices.com/ (accessed 15 May 2018).

Steinbach, C. M. (2001). “Markowitz Revisited: Mean-Variance Models in Financial Portfolio Analysis,” *SIAM Review* 43(1), 31-85.

Steiner, A. (2011). “Sharpe Ratio Contribution and Attribution Analysis,” *SSRN.* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1839166 (accessed 15 May 2018).