University of St. Gallen  
Graduate School of Business,  
Economics, Law and Social Sciences

RISK-ADJUSTED PERFORMANCE MEASURES – STATE OF THE ART

Master’s Thesis  
May 2010

Philipp Schmid*  
Supervised by  
Prof. Dr. Karl Frauendorfer

*ph.schmid@hotmail.ch, Unterdorfstrasse 2, 5505 Brunegg, Switzerland
Abstract

This thesis provides a comprehensive overview of the various applications of risk-adjusted performance measures (RAPMs) in the financial industry. RAPMs are used for efficient asset allocation, performance evaluation as well as for decisions on capital allocation within financial institutions.

Financial institutions face two major allocation problems: First, funds have to be invested with the objective to maximize the investor’s expected utility. Second, the risk capital of a financial institution must be optimally allocated on the different risky business activities. In the existing financial literature, these two research areas are generally treated separately, even though they rely on similar mathematical concepts. This thesis bridges the gap between the two areas and tests the most popular RAPMs with empirical data. It also discusses the potential manipulation of RAPMs and introduces spectral risk measures a potential future enhancement of existing RAPMs. The presented RAPMs include mean-variance performance measures, CAPM performance measures, downside performance measures and preference-based performance measures.

The findings suggest that for the investment selection process, the Sharpe ratio (SR) remains the leading RAPM, despite the drawback of not taking higher moments of a return distribution into account. Risk-adjusted return on capital (RAROC) is the preferred measure allocating the risk capital of financial institutions on their risky business activities.
Contents

1. Introduction ............................................................................................................................................... 1

2. Risk and Risk Management ........................................................................................................................... 3
   2.1 Definition of Risk ........................................................................................................................................ 3
   2.2 Financial Risk ........................................................................................................................................... 4
      2.2.1 Market Risk ......................................................................................................................................... 4
      2.2.2 Credit Risk ........................................................................................................................................ 5
      2.2.3 Operational Risk ............................................................................................................................ 6
      2.2.4 Model Risk and Liquidity Risk ....................................................................................................... 6
   2.3 Risk Management ...................................................................................................................................... 7
      2.3.1 Why Manage Financial Risk? ........................................................................................................ 8
   2.4 Risk Measurement ................................................................................................................................... 9
      2.4.1 Risk Measures ................................................................................................................................. 9
      2.4.2 Approaches of Risk Measurement ........................................................................................... 10

3. Risk-adjusted Performance Measures .......................................................................................................... 12
   3.1 Introduction ......................................................................................................................................... 12
   3.2 Mean-Variance Performance Measures ................................................................................................. 13
      3.2.1 The Sharpe Ratio ............................................................................................................................ 14
      3.2.2 Information Ratio ............................................................................................................................ 16
      3.2.3 The M² Measure .............................................................................................................................. 17
   3.3 CAPM Performance Measures ................................................................................................................ 18
      3.3.1 Treynor Ratio ................................................................................................................................. 19
      3.3.2 Jensen's Alpha and the Appraisal Ratio ......................................................................................... 19
      3.3.3 CAPM Performance Measures as traditional RAPMs .......................................................... 20
   3.4 Downside Risk Performance Measures ................................................................................................. 21
      3.4.1 The Modified Sharpe Ratio ............................................................................................................. 22
      3.4.2 The Sortino Ratio ........................................................................................................................... 23
      3.4.3 Omega and Sharpe-Omega ............................................................................................................. 24
      3.4.4 The Kappa Indices .......................................................................................................................... 27
      3.4.5 The Upside Potential Ratio ............................................................................................................ 29
      3.4.6 RAPMs based on Maximum Drawdown ................................................................................... 30
# List of Figures

<table>
<thead>
<tr>
<th>Number</th>
<th>Figure Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>CAPM and the Security Market Line</td>
<td>18</td>
</tr>
<tr>
<td>3.2</td>
<td>Simulation of Non-Normal Returns</td>
<td>22</td>
</tr>
<tr>
<td>3.3</td>
<td>Components of the Omega Ratio</td>
<td>25</td>
</tr>
<tr>
<td>3.4</td>
<td>Omega and Sharpe-Omega</td>
<td>26</td>
</tr>
<tr>
<td>3.5</td>
<td>The Kappa Indices</td>
<td>28</td>
</tr>
<tr>
<td>3.6</td>
<td>The Idea of AIRAP (CE) under CRRA</td>
<td>33</td>
</tr>
<tr>
<td>3.7</td>
<td>Positively Skewed Return Distributions</td>
<td>36</td>
</tr>
<tr>
<td>3.8</td>
<td>Power Utility Functions under CRRA 4 and 10</td>
<td>37</td>
</tr>
<tr>
<td>4.1</td>
<td>Economic Capital for Credit Risk</td>
<td>42</td>
</tr>
<tr>
<td>4.2</td>
<td>Value-at-Risk Calculation for Credit Risk</td>
<td>45</td>
</tr>
<tr>
<td>5.1</td>
<td>Line Chart &amp; Histogram</td>
<td>52</td>
</tr>
<tr>
<td>6.1</td>
<td>Coherent and a Non–Coherent Risk Aversion Function</td>
<td>55</td>
</tr>
<tr>
<td>A.1</td>
<td>Rankings based on Kappa Indices</td>
<td>65</td>
</tr>
</tbody>
</table>
List of Tables

3.1 Correlated Investment Opportunities .................................................................................................... 19
3.2 Unreliability of CAPM Decision Rules .................................................................................................. 26
3.3 Illustration of the MPPM CRRA = 4 ......................................................................................................... 41
3.4 Illustration of the MPPM CRRA = 10 ...................................................................................................... 41
3.5 Two Uncertain Investments ....................................................................................................................... 43
5.1 Main Characteristics of the Samples ....................................................................................................... 51
5.2 Ranking of SR, GSR, MPPM and Omega ................................................................................................. 52
5.3 Ranking of Kappa and Omega .................................................................................................................. 53
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIRAP</td>
<td>Alternative Investments Risk-Adjusted Performance</td>
</tr>
<tr>
<td>AMA</td>
<td>Advanced Measurement Approach</td>
</tr>
<tr>
<td>AR</td>
<td>Appraisal Ratio</td>
</tr>
<tr>
<td>CARA</td>
<td>Constant Absolute Risk Aversion</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Density Function</td>
</tr>
<tr>
<td>CE</td>
<td>Certainty Equivalent</td>
</tr>
<tr>
<td>CER</td>
<td>Certain Equivalent (excess) Return</td>
</tr>
<tr>
<td>CRRA</td>
<td>Constant Relative Risk Aversion</td>
</tr>
<tr>
<td>CVaR</td>
<td>Conditional Value-at-Risk</td>
</tr>
<tr>
<td>EC</td>
<td>Economic Capital</td>
</tr>
<tr>
<td>ES</td>
<td>Expected Shortfall</td>
</tr>
<tr>
<td>EVT</td>
<td>Extreme Value Theory</td>
</tr>
<tr>
<td>GSR</td>
<td>Generalized Sharpe Ratio</td>
</tr>
<tr>
<td>HPM</td>
<td>Higher Partial Moment</td>
</tr>
<tr>
<td>i.i.d.</td>
<td>independent and identically distributed</td>
</tr>
<tr>
<td>IR</td>
<td>Information Ratio</td>
</tr>
<tr>
<td>IRB</td>
<td>Internal Rating Based</td>
</tr>
<tr>
<td>LPMs</td>
<td>Lower Partial Moments</td>
</tr>
<tr>
<td>MAR</td>
<td>Minimum Acceptable Return</td>
</tr>
<tr>
<td>MPPM</td>
<td>Manipulation Proof Performance Measure</td>
</tr>
<tr>
<td>MSR</td>
<td>Modified Sharpe Ratio</td>
</tr>
<tr>
<td>MVaR</td>
<td>Modified Value-at-Risk</td>
</tr>
<tr>
<td>P&amp;L</td>
<td>Profit and Loss</td>
</tr>
<tr>
<td>RAPM</td>
<td>Risk-Adjusted Performance Measure</td>
</tr>
<tr>
<td>RAROA</td>
<td>Risk-Adjusted Return on Assets</td>
</tr>
<tr>
<td>RAROC</td>
<td>Risk-Adjusted Return on Capital</td>
</tr>
<tr>
<td>RORAA</td>
<td>Return on Risk-Adjusted Assets</td>
</tr>
<tr>
<td>SML</td>
<td>Security Market Line</td>
</tr>
<tr>
<td>SoR</td>
<td>Sortino Ratio</td>
</tr>
<tr>
<td>SR</td>
<td>Sharpe Ratio</td>
</tr>
<tr>
<td>TR</td>
<td>Treynor Ratio</td>
</tr>
<tr>
<td>UPR</td>
<td>Upside Potential Ratio</td>
</tr>
<tr>
<td>VaR</td>
<td>Value-at-Risk</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

In financial institutions there are various reasons to apply risk-adjusted performance measures (RAPMs), either ex-ante for asset and capital allocation decisions or ex-post for evaluating the actual performance of allocation decisions.

Senior risk managers and asset managers apply RAPMs for distinct purposes. While senior risk managers use RAPMs to allocate risk capital across an institution’s risky business activities in an efficient manner, asset managers apply RAPMs in order to provide their investors with the best possible return to risk ratio to their investors. In either area academics and practitioners have developed numerous RAPMs. The development in the area of capital allocation process was mainly driven by regulator. In the area of investment selection process this development was favored by hedge funds and other alternative investments. Today, the most widely accepted RAPMs are the Sharpe ratio (SR) for asset allocation and the risk-adjusted return on capital (RAROC) for capital allocation within financial institutions.

From this basis, this thesis provides a comprehensive overview of the major RAPMs that are applied in the financial industry. In contrast to the existing financial literature, which generally treats these two areas separately, both areas – the area of optimal capital allocation within an institution and the area of optimal asset allocation – are discussed.

As RAPM is a generic term used to describe all techniques used to adjust returns for the risks incurred in generating those returns, the thesis focuses in particular on the risk component. The return element usually poses few major problems. Major advantages and disadvantages of the different RAPMs are presented and where the application of certain RAPM is accurate and where other RAPMs might be better suited is investigated. Potential manipulations of existing RAPMs are also discussed in detail. In the empirical section, the different RAPMs are compared subject to their properties for measuring return distributions with different distribution characteristics.
In an outlook a potential enhancement of existing RAPMs by the application of spectral risk measures in the denominator of the RAPM ratio is presented. Such a performance measure has the potential to further improve the accuracy of risk-adjusted performance measurement.

The thesis is structured as follows. Chapter 2 provides a basic definition of risk and an introduction to the duties of modern risk management in financial institutions. Chapter 3 introduces the major RAPMs applied by asset managers and investors in the investment selection process. Chapter 4 is dedicated to the RAPMs used by the senior management of a company to allocate an institution’s risk capital in an efficient manner. Chapter 5 examines six investment strategies by applying the introduced RAPMs to empirical data. Finally, chapter 6 summarizes the main conclusions and completes the thesis with an outlook of spectral risk measures as a potential future enhancement of existing RAPMs.
Chapter 2

Risk and Risk Management

2.1 Definition of Risk

Risk can be defined in various ways. The Oxford English Dictionary (2005) for instance defines risk as “the possibility of something bad happening at some time in the future, a situation that could be dangerous or have a bad result”. Obviously, that this definition only mentions the drawback, i.e. the downside of risk, whereas the potential for a gain, i.e. the possible upside, is neglected. It is not surprising that therefore in normal linguistic usage people usually associate risk only with negative events such as car accidents or environmental disasters. Yet, while risk might harm those who are exposed to it, it however also offers benefits for those who succeed in using it to their advantage. Therefore, in a formal financial context the term “risk” implies in addition to the negative element, a positive one. Commonly, in order to exploit a market opportunity or rather an uncertain future return, you have to take risk and sacrifice current resources. A definition such as the one from DeLoach (2000) who defines risk “as the distribution of possible outcomes in a firm’s performance over a given time horizon due to changes in key underlying variables” (p. 66) might in this context be more accurate.

Crouhy, Galai and Mark (2005) emphasize that the natural human understanding of risk is fairly sophisticated. As a matter of fact, people usually budget their expected daily life costs, and even if these costs are large, they are usually not seen a threat as they are reasonably predictable. The real risk arises if those costs suddenly soar in a completely unforeseen way, or if costs appear out of nowhere and consume the money which had been saved for these expected outlays. Therefore, people relate risk in general to unexpected losses (unexpected costs) which exceed their expected losses (expected costs). This distinction between unexpected and expected costs, is essential in modern risk management concepts such as the allocation of economic capital and risk-adjusted pricing.
2.2 Financial Risk

Entrepreneurial activities and risk-taking are inextricably linked to each other. Risk-taking is an essential component of doing business considering basically every entrepreneurial activity is exposed to a greater or lesser degree of uncertainty. One can think of risk as the uncertainty about the future demand for products and services, changes in the business environment and competition and production technologies. In addition to these general business risks, there also exist risks that are caused by the capital structure of a company such as market risks, credit risks, operational risks and liquidity risks.

Risk is discussed in the context of banks and other financial institutions, consequently the main risk types encountered in the financial industry will be introduced in the next section. Following the regulatory approach in the global banking industry, the three major risk categories are market risk, credit risk as well as operational risk. Nevertheless they do not form an exhaustive list of possible risks affecting a financial institution, as various other risks such as reputation risk, strategic risk, liquidity risk and model risk may occur. Particularly, the latter two (i.e. liquidity risk and model risk) have received a lot of attention recently and thus will be briefly discussed as well.

2.2.1 Market Risk

According to McNeil, Frey and Embrechts (2005) the best known type of risk in banking is market risk, which is the risk of change in the value of a financial security (e.g. a derivative instrument) due to changes in the value of their underlyings, such as stock prices, bond prices, exchange rates and commodity prices. In other words, it is risk that changes in financial market prices and rates, which will reduce the value of a security or a portfolio. Market risk usually arises from both unhedged positions as well as imperfect hedged. Crouhy et al. (2005) distinguish four major types of market risks:

- **Interest-Rate Risk** is caused by changes in the market interest rate. Usually the value of fixed-income securities such as bonds, is highly dependent on those interest rates. For instance, when market interest rates rise, the value of owning an instrument offering fixed interests payments falls. Moreover, Hull (2007) emphasizes that managing interest-rate risk is more complex than managing the risk arising from other market variables such as equity prices, exchange rates and commodity prices. On account of the many different interest rates in a given currency, e.g. treasury rates, interbank borrowing and lending rates, mortgage rates etc. These tend to move together, but are normally not perfectly correlated. Furthermore the term structure is only known with certainty for a few specific maturity dates, while the other maturities must be calculated by interpolation.

- **Equity-Price Risk** is associated with the volatility of stock prices. The general market risk of equity refers to the sensitivity of the value of a security to change in the market portfolio. According to the portfolio theory, the market risk, i.e. the systematic risk, cannot be
eliminated through portfolio diversification, whereas the unsystematic risk can be completely diversified away.

- **Foreign-Exchange Risk** arises from open or imperfectly hedged positions in a particular foreign currency. These positions may arise due to natural consequences of business operations such as cross-border investments. The major drivers of foreign-exchange risk are imperfect correlations in the movement of currency prices and fluctuations in international interest rates. Therefore, one of the major risk factors large multinational corporations are exposed to, are foreign exchange volatilities, which may on the one hand diminish returns from cross-border investments or on the other hand increase them.

- **Commodity-Price Risk** differs considerably from interest-rate and foreign-exchange risk, as commodities are usually traded in markets where the supply of most commodities lies in the hands of a just few market participants, which may result in liquidity issues often followed by exacerbating high levels of price volatility. Moreover, storage costs heavily affect commodity prices which vary considerably across commodity markets (e.g. from gold, to electricity, to wheat) on the one hand and on the other hand the benefit of having a certain commodity on stock provides a convenience yield.

### 2.2.2 Credit Risk

Another important risk category is credit risk: The risk that a change in the creditworthiness of a counterparty affects the value of a security or a portfolio. Not receiving all promised repayments on outstanding investments such as loans and bonds due to default of the debtor, are the extreme cases. When a company goes bankrupt, the counterparty usually loses the part of the market value that cannot be recovered following the insolvency. The amount expected to be lost is normally called the loss given default whereas the recovery rate is defined as the market value immediately after default (see Hull, 2007). A change in the creditworthiness usually does not imply a default, but rather that the probability of a default increases. A deterioration of the credit rating leads to a loss for the creditor since a higher marked yield is required to compensate for the increased risk which results in a value decline of the debts (e.g. bonds). Crouhy et al. (2005) stressed that institutions are also exposed to the risk that a counterparty might be downgraded by a rating agency. Rating agencies such as Moody’s and Standard & Poor (S&P) provide ratings that describe the creditworthiness of corporate bonds and therefore provide information about default probabilities. If a company is downgraded by a rating agency due to a negative long-term change in the company’s creditworthiness, the value of the counterparty’s securities diminishes.
2.2.3 Operational Risk

A further important risk category recently receiving a lot of attention is operational risk. Operational risk is not only more complex to quantify than market and credit risk but also more difficult to manage as it is a necessary part of doing business.

Hull (2007) mentions that there are many different definitions to operational risk and that it is tempting to consider it as a residual risk category, covering any risk faced by a bank that is not either market or credit risk. Nevertheless, this definition of operational risk might be too broad. To define it straightforward, as its name implies, it is the risk arising from operations. Thus, the risk relates to potential losses resulting from inadequate systems, management failures, faulty controls, frauds, and human errors.

According to the Basel Committee on Banking Supervision (2004) operational risk is defined “as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events” (p. 137). Apparently the regulator includes, besides the impact of internal risks, the impact of external risks such as natural disasters (e.g. earthquakes and fires).

Operational risk is not independent from other financial risks. Operational risk losses are for instance frequently contingent on market movements, which enhance the complexity of their classification. One can relate it to a trader taking huge risk in order to receive a tremendous bonus at the end of the year. If – as a result of adverse market movements – the bank suffers huge losses, the risk that led to it can be classified as either operational or market risk, depending on whether the trader was allowed to take that much risk or not.

2.2.4 Model Risk and Liquidity Risk

While banks have always been exposed to threats such as bank robberies and white-collar frauds, one of today’s most serious threats is caused by the valuation of complex derivative products, which has come to be known as model risk.

Since Black, Scholes and Merton in 1973 published their famous option-pricing model, there has been a tremendous increase in the complexity of valuation theories. These models allow for a pricing of a huge number of financial innovations such as caps, floors, swaptions, credit derivatives, and other exotic products. As a negative side effect to the rise in complexity of financial products, the accompanying model risk has increased as well. For instance, Derman (2004) emphasizes in his book “My Life as a Quant” that this increase was essentially caused by the nature of the models used in finance. In principle, most of these applied models, including the Black-Scholes option-pricing model, have been derived from models encountered in physics. While models of physics are highly accurate, models of finance describe the behavior of market variables which in turn unlike in physics depend on the actions of human beings. Therefore, the models are at best approximate descriptions of the market variables. As a result the use of such models in finance is always accompanied – to a greater or lesser extent – by model risk.
Hull (2007) mentions two main types of model risk. The first type concerns the risk that a valuation model could provide wrong prices, which can lead to an investor to buy or sell a product at a price that is either too high or too low. The second type relates to models that are used to assess risk exposure and to derive an appropriate hedging strategy in order to mitigate losses. For instance, a company may use a wrong or inadequate model to hedge its positions against an adverse movement of the underlying assets.

It, however, is important to bear in mind, that a theoretical valuation model is only essential for pricing products that are relatively or even completely illiquid. If there is an active market for a product, market prices are usually the best indicator of an asset's value and therefore pricing models only play a minor role.

The risk that a firm does not have enough cash and cash equivalents in order to meet its financial obligations as well as the risk of not having enough buyers or sellers on the market is known as liquidity risk. Crouhy et al. (2005) distinguished two dimensions of liquidity risk, namely funding liquidity risk and asset liquidity risk. Funding liquidity risk relates to a firm's ability to raise the required cash to meet its liabilities. Asset liquidity risk, on the other hand, arises if an institution cannot execute a transaction at the prevailing market price due respectively to a lack of supply and demand.

### 2.3 Risk Management

It is beyond dispute that the future cannot be exactly predicted, as it is always uncertain to a certain degree. However, the risk that is caused by this uncertainty can be managed. Risk management is therefore how financial institutions actively select the overall level of risk that, given their risk-taking ability, is optimal for them. Yet it is important to note that risk management also encompasses the duality of the term risk, as risk management is not only about risk reduction. According to McNeil et al. (2005), a bank's attitude to risk is rather active than defensive, as banker actively and willingly take on risk in order to benefit from return opportunities. Risk management can therefore be seen as the core competence of a bank. Bankers are using their expertise, market position and capital structure to manage risks by restructuring and transferring them to various market participants.

Crouhy et al. (2005) on one hand refer risk management to be widely acknowledged as one of the most creative forces in the world's financial markets. An example, is the rapid development of the huge market for credit derivatives, which emphasize the dispersion of risk (i.e. the credit risk exposure) of an institution to those who are willing, and presumably able, to bear it.

On the other hand, Crouhy et al. (2005) mention extraordinary failures in risk management such as Long-Term Capital Management and the string of financial scandals associated with the millennial boom in equity and technology markets (e.g. Enron and WorldCom). These are only a few examples
of where risk management has not been able to prevent market disruptions and business accounting scandals.

The reason for this ambiguity lies in the ambivalent nature of the new techniques in risk management. They enhance market liquidity leading to a far more flexible, efficient and resilient financial system. At the same time, however, they are according to Instefjord (2005) also a potential threat to bank stability and may expose a financial institution to even more risk.

Today’s risk management has changed compared to traditional risk management, which was basically identifying, measuring, managing, and minimizing risk. The role of today’s risk management has changed from minimizing risk to efficient capital allocation and become more important, as it can increase business profitability by allocating capital and the entrepreneurial attention on the areas with the highest risk and return ratio. Hence, the application of RAPMs has become popular in the finance industry in order to evaluate and compare different business units.

### 2.3.1 Why Manage Financial Risk?

An important issue is whether there should be any investment in risk management in the first place. Assuming frictionless markets, in equilibrium all risks should be appropriately priced. Hence, if there were no capital market imperfections, Modigliani and Miller’s Proposition I – the so-called capital structure irrelevance theorem – would apply and the problem of capital allocation would be nonexistent. As a result there would be no reason of why a financial institution would want to manage risk at all. Yet, financial markets are neither frictionless nor are they always in equilibrium. As an example, a fundamental role of banks and other financial institutions is to invest in illiquid financial assets (e.g. loans to small or medium sized companies). These assets cannot be traded frictionless in the capital markets, due to their information intensive nature. In fact, financial institutions and banks, in particular, face market imperfections such as costs of financial distress, transactions costs and regulatory constraints, with the consequence that risk management, capital structure and capital budgeting are interdependent (see, for instance, Copeland, 2005). Consequently, there indeed exist various reasons in reality for managing risk. As stated by McNeil et al. (2005) most stakeholders, including shareholders, management and regulators, have an incentive in the management of risk, since it is usually beneficial for a financial institution. Modern society relies on a smooth functioning of the financial system. It is therefore common in best interest to regulate and manage the risk imperiling such systems in order to avoid systemic risk, which in extreme situations may disrupt the normal functioning of the entire financial system. The literature provides various other examples which are in favor of investments in risk management, such as it reduces the costs of financial-distress and also the costs of taxes. Reader interested in a more comprehensive overview may refer Froot and Stein (1995).
2.4 Risk Measurement

2.4.1 Risk Measures

A central issue in modern risk management is measuring and quantifying risk. To set risk limits as well as determining adequate risk capital as a cushion a financial institution requires against unexpected future losses, belong to the most important functions of risk measurement.

Various methods exist to measure these risks, all with the target of capturing the variation of a company’s performance. Bessis (2002) distinguishes three categories of risk measures.

- **Volatility** captures the standard deviation of a target variable around its mean. The standard deviation is the square root of the average squared deviation of a target variable from its expected value. Since volatility captures both upside and downside variations, it is a symmetric risk measure which assigns the same amount of risk to deviations above and below the mean. Therefore, volatility lacks in providing a complete picture of risk in the case the target variable has an asymmetric distribution.

- **Sensitivity** captures the deviation of a target variable due to a movement of a single underlying parameter. Sensitivities are normally market risk related as they relate value changes to market parameters such as interest-rate risk. Among all sensitivity measures, the most famous ones are the Duration for bond portfolios and the Greeks for portfolios of derivative instruments. Even though these measures provide useful information regarding the robustness of a portfolio with respect to certain events, they fail to quantify the overall riskiness of a position. Furthermore, they cause problems when risks need to be aggregated (see McNeil et al., 2005).

- **Downside Risk Measures** are – unlike the volatility – asymmetric risk measures which focus on adverse deviations of a target variable only. The lower partial moments (LPMs) of order k and the quantile-risk measures such as the Value-at-Risk (VaR) and the expected shortfall (ES) are the most widely used downside risk measures, VaR being the most prominent one. These downside risk measures focus exclusively on extreme downside moves of the risk factors, rather than considering both upside gains and downside losses. This makes downside risk measures intuitively the most reasonable risk measure, as they are consistent with the human natural asymmetric perception of risk. Measures based on the concept of downside risk are useful in particular when the target variable has a highly skewed distribution, given that skewed distributions need more than the first two statistic moments to be adequately specified. However, if the distribution of a variable is symmetric and not asymmetric, downside risk measures do not provide a more comprehensive picture than the symmetric volatility measure. Unfortunately, the calculation of most downside risk measures is fairly complex, especially when considering derivative financial products with asymmetric payoffs.
Already Markowitz (1959) recognized the limitations of the mean-variance approach and suggested to use downside risk measures rather than the volatility measure. Recent risk management literature has focused on downside risk measures such as the VaR, whereas average risk measures, in particular the volatility measure, play a minor role (Martellini, Priaulet & Priaulet, 2003). Intuitively this makes sense, as in risk management it is usually most important to obtain a feeling of what deteriorating a financial situation can become in the case certain risk factors turn out to be adverse.

2.4.2 Approaches of Risk Measurement

In order to provide a comprehensive overview of this subject, it is useful to refer to a slightly different approach mentioned by McNeil et al. (2005), which give an overview of existing techniques to measure risk in financial institutions. Moreover, these approaches are grouped into four different categories:

- **The Notional-Amount Approach** is the oldest approach quantifying the risk of a portfolio of risky assets. The calculation of the risk is fairly simple and the sum up of the notional values are weighted by each security’s risk factor class. However, even though this approach seems to be crude, McNeil et al. (2005) mention that some “variants of this approach are still in use in the standardized approach of the Basel Committee rules on banking regulation” (p. 35).

- **Factor-Sensitivity Measures** are an approach identical to the risk measure category sensitivity mentioned above. A further explanation is therefore not necessary (see section 2.4.1, Sensitivity).

- **Risk Measures Based on a Loss Distribution** are the most popular approach, being that most modern risk measures are based on a profit and loss (P&L) distribution. A P&L distribution tries to provide an accurate picture of the existing risk in a portfolio or even of the financial institution’s overall position in risky assets. The P&L distribution is the distribution of the change in value $V(s+\Delta)-V(s)$, where $V(s)$ is the value of a portfolio at time $s$ and $\Delta$ is a given time horizon such as 1 or 10 days. Since the focus is on the probability of the occurrence of large losses or more formal the upper tail of the loss distribution, it is according to McNeil et al. (2005) common to drop the P from P&L and to simply use the term loss distribution. Both variance and VaR are based on such a loss distribution and accordingly rely on historic data.

- **Scenario-Based Risk Measures** are a rather new approach to measure the risk of a portfolio, even though it actually pre-dates VaR modeling approach. As a matter of fact, the first commercial application of scenario stress testing was already established in the 1980s with the Chicago Mercantile Exchange to determine its margin requirements. The risk of a portfolio is measured by considering possible future scenarios (i.e. risk-factor changes) such as a rise in the exchange rate and a simultaneous drop in an underlying stock. The total
portfolio risk is then defined as the maximum loss of the portfolio taking all scenarios into consideration. This corresponds more or less to a sensitivity analysis that examines the loss profile of a portfolio, by considering a number of changes in certain risk factors. Given the tremendous number of possible historical and hypothetical scenarios, it is important to distinguish between the major risk drivers of a portfolio and the minor ones. Commonly, these major risk factors are based on the market risk since these risk factors are relatively easy to obtain, especially as compared with credit risk and operational risk.

Today, loss distributions are the most popular approach to quantify risk. Yet, when working with loss distributions, two major problems emerge. First, loss distributions are based on historical asset returns. This historical data might be of limited use in predicting future risks. Second, it is difficult to accurately estimate loss distributions; in particular for large portfolios whereas their calculation becomes extremely complex. Nevertheless, these issues are according to McNeil et al. (2005) not arguments against the use of loss distributions. Rather, it is important to improve the way these loss distributions are estimated and to use more caution when applying risk measures based on loss distributions.

Besides the approaches presented above, another approach, the Extreme Value Theory (EVT) has received a lot of attention recently. EVT provides a framework to formalize the study of behavior in the tails of a distribution. Similar to the scenario stress tests, EVT tries to capture extreme events (also referred to as low probability events) that according to the loss distribution have a probability of virtually zero percent. For instance, a move of five standard deviations in a market variable is such a rare event that under the assumption of normally distribution this should occur only once every 7’000 years. Yet, they actually do occur from time to time. Best example is the subprime crisis that began in mid-2007, revealing that the current regulatory capital framework for banks does not capture some key risks. Moreover, the crisis showed that a quantile-based estimation of risk capital usually cannot cover the extreme losses that can incur in unexpected exceptional circumstances. As a result, new approaches have been developed in the last years that look beyond volatility and VaR. (Alexander, 2008b; Haan and Ferreira, 2006)
Chapter 3

Risk-adjusted Performance Measures

3.1 Introduction

Prior to the development of risk-adjusted performance measures (RAPMs), the performance of individual investments and portfolios of investments was solely assessed by comparing total returns. This approach of performance evaluation was subsequently supplemented by comparing the total return with an unmanaged benchmark, such as the S&P 500. Although these benchmarks became over the years more accurate, reflecting more precisely on the composition of a relevant portfolio, the profit was still measured on a total return basis.

However, total return is an incomplete performance measure, as it entirely ignores the factor of risk. This is an obvious drawback since it is well known that a portfolio can increase expected total return by accepting a higher level of risks. Consequently, along with the introduction of the modern portfolio theory, practitioners and academics developed absolute performance indicators, also referred to as RAPMs, with the objective of adequately adjusting returns for the risk incurred in generating those returns.

The focal point of this chapter is to provide a comprehensive overview over the most widely used RAPMs for the investment selection process. Unfortunately, the literature concerning RAPMs for the investment selection process is over abundant and confusing, owing to the suggestion of a variety of alternative solutions for the same problem (see Dowd, 2001). An appropriate categorization is therefore crucial to better understand the different properties of these RAPMs.

In order to account for this the RAPMs have been divided into four categories, namely, mean-variance performance measures, CAPM performance measures, downside risk performance measures and preference-based performance measures. This categorization is convenient as all RAPMs within the same category are based on the same assumptions and therefore also share most of the same advantages and disadvantages. As stressed by Ornelas et al. (2010) and Koekebakker and Zakamouline (2009), the choice of one RAPM over another RAPM does affect the performance
evaluation (i.e. the ranking of investment opportunities). Hence, if investors apply RAPMs for assessing risky investment opportunities, they should be especially aware of the shortcomings of the applied RAPM, as some might be appropriate in a particular case, but rather inappropriate in another case.

When using RAPMs for assessing risky investment opportunities ex-ante or for evaluating the actual performance of investments ex-post, it is moreover important to bear in mind that RAPMs are usually based on historical data. An accurate database is therefore crucial, since an inaccurate database might have a significant impact on the applied performance measures. While Liang (2003) enumerates factors that affect the quality of the database, the database can also suffer from several biases such as the survivorship bias, the instant history bias and the selection bias, to name but a few (Géhin, 2004).

Another problem in the database is caused by the presence of either positive or negative autocorrelation. The sample estimates are usually based on monthly, weekly or even daily data, whereas the quantities used for most RAPMs are quoted in annualized terms. The annualization is normally performed under the assumption that returns are independent and identically distributed (i.i.d.). This, however, will not hold if the returns of the investment are autocorrelated. In concrete terms, a positive autocorrelation will lead to a higher standard deviation, which in turn will reduce the value of the applied RAPM. Lo (2002) illustrated that the adjustment for autocorrelation, indeed, has a highly significant impact on the applied RAPM. However, for this thesis it is sufficient enough to just be aware of these issues, as it is beyond the scope of this thesis to go into further detail.

### 3.2 Mean-Variance Performance Measures

The mean-variance framework developed by Markowitz (1959) is the traditional approach to model the trade-off between risk and reward. It was built on the concept of the utility function introduced by Von Neumann and Morgenstern (1947). The framework formalizes how risk-averse investors should select between various risky assets, given certain idealized assumptions, in order to maximize their expected utility.

In this setting, the reward is represented by the expected return, whereas the risk is defined as the variance of the returns. Assuming that the investor has an exponential utility function and that returns are normally distributed, one can find the optimal investment decision by applying the

---

1 This stands in contrast with the findings of Eling and Schuhmacher (2007), who conclude that most RAPMs provide virtually identical rankings. This, however, seems rather implausible as stressed by Ornelas et al. (2010).

2 See for instance Verbeek (2008) Chapter 4 for a detailed explanation of how to measure and adjust for autocorrelation.
mean-variance criterion. In other words, an investor’s utility function can be approximated by a function of the mean and the variance of the return distribution. As a result, when using the variance around the mean as the risk measure, investors are assumed to have no preference for upside or downside risks. This is, however, most certainly not the case in reality as investors do have a preference for upside risk and an aversion to downside risk. Moreover, decisions based on the mean-variance framework take only the first two moments of a distribution into account, while potential distinctions among investments in higher moments are neglected (i.e. skewness and kurtosis). RAPMs based on the mean-variance framework are therefore only meaningful measures of performance as long as risk can be adequately measured by the variance. This is the case when returns are normally or at least approximately normally distributed.

### 3.2.1 The Sharpe Ratio

The basic Sharpe ratio (SR) was introduced by Sharpe (1966) over 40 years ago and is – despite its age – even today one of the most popular and commonly used RAPMs. The financial industry applies the SR in many different contexts, from performance attribution to tests of market efficiency to risk management. A recent survey from Goltz and Schroeder (2008) even shows that more than 67% of hedge funds managers construct their portfolio based on the Sharpe optimization, which is fairly surprising in light of the many shortcomings of the SR in evaluating hedge funds (see section 3.4). The Sharpe rule proposes to select the asset with the highest expected SR (i.e. the highest risk-return trade-off), with the SR being calculated as the excess of the expected return of an investment over the risk-free rate divided by the standard deviation of the investment’s return distribution, i.e.

$$ SR_i = \frac{E(R_i) - R_f}{\sigma_i}, $$

where $E(R_i)$ is the expected return of investment $i$, $R_f$ is the risk-free rate and $\sigma_i$ is the standard deviation of the return of investment $i$.

Intuitively, the SR can be interpreted as an investment’s excess return per unit of risk. Hence, the SR is commonly also interpreted as a reward-to-risk ratio, where both risk and return are captured in a single measure. A rising expected return and a falling standard deviation are both positive events from an investor perspective, leading to a rise in the SR. The SR is used to characterize how well the return of an investment compensates the investor for the risk taken: The higher the SR the better the risk-adjusted performance.

The SR provides sufficient information to choose between two investment opportunities ex-ante or to evaluate alternative investment strategies ex-post as long as the investments are uncorrelated.

---

3 It should be noted that the mean-variance framework is sometimes mistakenly justified by the assumptions that either investors have a quadratic utility function or that returns are normally distributed. Yet as for instance stressed by Pézier (2008), the condition that investors have a quadratic utility function is not consistent with the behavior of a rational investor because it does not satisfy the principle of non-satiation and also implies increasing absolute risk aversion (IARA) and is thus incorrect.
Risk-adjusted Performance Measures

with existing positions in a portfolio. It was Sharpe (1994) himself who acknowledged that the SR may not give correct answers if one or more investments are correlated with existing positions in a portfolio. It is therefore possible that even if investment A has a higher individual SR than investment B, investment B might be superior to investment A if the correlation effect is taken into account (see table 3.1). This implies that the individual SR cannot generally be relied upon to give the correct answer. Therefore, the SR should be applied to the alternative portfolios of investments – which are equal to each investment plus the existing positions – rather than to the alternative investment opportunities alone.

### Table 3.1

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>Risk-free asset</th>
<th>Existing position</th>
<th>Investment plus existing position</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expected return</strong></td>
<td>10.0%</td>
<td>8.0%</td>
<td>5.0%</td>
<td>9.0%</td>
<td>9.5%</td>
</tr>
<tr>
<td><strong>Standard deviation of return</strong></td>
<td>10.0%</td>
<td>8.0%</td>
<td>0.0%</td>
<td>9.0%</td>
<td>8.2%</td>
</tr>
<tr>
<td><strong>Expected excess return</strong></td>
<td>5.0%</td>
<td>3.0%</td>
<td>-</td>
<td>4.0%</td>
<td>4.5%</td>
</tr>
<tr>
<td><strong>Correlation with existing portfolio</strong></td>
<td>0.50</td>
<td>-0.50</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Sharpe Ratio (SR)</strong></td>
<td>0.50</td>
<td>0.375</td>
<td>0.0</td>
<td>0.547</td>
<td>0.819</td>
</tr>
</tbody>
</table>

Similar to many other RAPMs the SR does not depend on the amount invested in the risky asset (i.e. is scale invariant), nor on the investor’s degree of risk aversion. In contrast, it is highly dependent on the investment horizon, a one-year horizon usually being the standard. According to Pézier (2008), the SR is only an accurate RAPM (i.e. maximizes the utility of an investor in a mean-variance framework) when the following six conditions are at least approximately met:

(A1) The investor’s only concern is total wealth at the end of an investment period, where he always prefers more wealth to less wealth.

(A2) The investor’s total wealth is optimally allocated to a risky investment and the risk-free asset, with both being available in unlimited positive or negative amounts (i.e. long or short positions).

(A3) The preferences of the investor are fully determined by the expected return and its variance and standard deviations, respectively.

(A4) The investor is risk averse in the sense that given equal expected returns, he always prefers the investment with the lowest variance (i.e. standard deviation).

(A5) The forecast of the risky investment returns is normally distributed with $N(\mu, \sigma^2)$. 
(A6) The risk attitude of the investor is described by a negative exponential utility function of wealth, i.e. $u(w) = -\exp(-\lambda w)$, where $\lambda > 0$ is the constant absolute risk aversion (CARA).

Since this is an example of a concave utility function, this implies a risk averse investor (see figure 3.6).

Yet conditions (A1) to (A6) are often not met in practice. For instance, many investors are willing to take more risk when becoming wealthier, which is in contrast to the fixed CARA in condition (A6). Moreover, unlike (A5), return forecasts are usually not normally distributed, and in contrast to (A3), investors are usually also averse to negative skewness (i.e. a left skewed distribution) and to positive excess kurtosis (i.e. heavy tails, implying extreme events are more likely than under normality). Solely condition (A4) appears to be generally consistent with the reality since investors are indeed risk averse.

Condition (A5) is the most important for the SR to provide an accurate performance evaluation. Therefore it is usually sufficient to assume that it approximately holds, when assessing alternative investment opportunities. However, as we will see in sections 3.4 and 3.6, returns are often not normally distributed and the SR is furthermore sensitive to abnormal shapes of the return distribution. Thus, the SR is only an accurate RAPM as long as returns are approximately normally distributed. Due to the fact that all mean-variance RAPMs rely on these assumptions, they are all subject to the same condition whereas at least (A5) must approximately hold.

### 3.2.2 Information Ratio

Over 25 years after Sharpe (1966) introduced the SR, Sharpe (1994) proposed a generalization of the original SR by relating performance to a benchmark instead of the risk-free rate. This new version of the SR is in some publications also referred to as the information ratio (IR). Apparently, the IR is equivalent to the original SR where the benchmark is a risk-free asset (e.g. government bonds).

The IR is defined as the ratio of the investment’s excess return over that of a benchmark (usually a well diversified index such as the S&P 500), divided by the standard deviation of the excess return (also referred to as the tracking error$^4$), i.e.

$$ IR_i = \frac{R_i - R_b}{\sigma_{i-R_b}} $$

where $R_i - R_b$ is the excess return and $\sigma_{i-R_b}$ is its standard deviation. The main idea behind the IR is that the benchmark serves as initial hypothetical investment, with the goal to allocate the money in a way that is superior to the benchmark in terms of risk and expected return. Therefore, a high expected IR implies that an investment or investment portfolio yields substantially higher expected

---

$^4$The tracking error is used as a measure of how closely an investment or portfolio follows the index to which it is benchmarked. Thus, the tracking error is a measure of the deviation from a benchmark.
return than the benchmark for relatively little extra risk. In contrast a low expected IR indicates little extra return relative to the higher risks taken (see Dowd, 2001).

### 3.2.3 The M² Measure

The M-squared measure, owing its name to the authors Modigliani and Modigliani (1997), focuses on assessing the performance of a portfolio compared to a benchmark (usually an index such as the S&P 500). They proposed a new RAPM, which intuitively is fairly easy to understand by investors. The main idea behind this measure is that an investment’s risk level is adjusted to match that of a benchmark. By leveraging or deleveraging the portfolio to match the benchmark’s risk exposure the returns can be easily compared.

This method has two major advantages: First, it reports the risk-adjusted performance of an investment as a percentage, which is easily understandable by investors; and second, investors know the degree of leverage that is required to attain a certain return. Therefore, risk-averse investors can use this information to reduce their risk by delivering their portfolio (i.e. selling part of the portfolio and buying risk-free securities), whereas aggressive investors are able to leverage their portfolio (i.e. borrowing money and investing in the portfolio). The M-squared can easily be computed by multiplying the SR with the benchmark’s standard deviation and then adding the risk-free rate, i.e.

\[ M_{i}^2 = \frac{R_i - R_f}{\sigma_i} \cdot \sigma_m + R_f, \]

where \((R_i - R_f)/\sigma_i\) is equal to the \(SR_i\) in (3.1), \(\sigma_m\) is the standard deviation of the benchmark and \(R_f\) is the risk-free rate.

The leverage factor can be calculated by dividing the standard deviation of the market by the standard deviation of the investment, i.e.

\[ L_i = \frac{\sigma_m}{\sigma_i}, \]

where \(\sigma_m\) is the standard deviation of the market and \(\sigma_i\) is the standard deviation of investment \(i\). Hence, if \(L_i > 1\) portfolio \(i\) is less risky than the benchmark, and if \(L_i < 1\) portfolio \(i\) is more risky than the benchmark. For instance, in order to reduce the risk of a volatile technology fund, Modigliani and Modigliani added risk-free Treasury bills to the portfolio until it corresponds with the index. In contrast, for a conservative fund, they leveraged the fund in order to match the risk profile of the index.

Rankings portfolios with the M-squared yields the same results as the rankings based on the SR. The only difference is that the M-squared expresses the score in basis points, which is easier to understand by the average investor.
3.3 CAPM Performance Measures

With the mean-variance framework Markowitz (1959) laid the foundation for the capital asset pricing model (CAPM), which was independently developed by Treynor (1965), Sharpe (1964) and Linter (1965).

Figure 3.1: In the CAPM equilibrium no single asset is supposed to have an abnormal return of alpha (i.e. $\alpha_i$) above (or below) the security market line (SML), which is formally defined as $SML = \beta_i (E(R_i) - R_f) + R_f$. Hence, if the market is not in the equilibrium, an investor should purchase any asset that yields a positive alpha.

The CAPM is based on two general assumptions: First, investors are assumed to choose portfolios according to the mean-variance framework. Second, the CAPM assumes that there is only one source of risk for which investors are being rewarded\(^5\). The risk is referred to as the systematic risk, which cannot be diversified away by holding a large portfolio of different risky assets. In the market equilibrium, the expected return of any single asset is proportional to the expected excess return on the market portfolio, i.e.

$$E(R_i) - R_f = \beta_i (E(R_m) - R_f),$$

where the left-hand side is the expected excess return of asset $i$ and the right-hand side is the CAPM equilibrium. $\beta_i$ represents the coefficient for the systematic risk and therefore equals the sensitivity of the asset return to changes in the market return. In other words, the return required on any asset is equal to the risk-free rate of return plus a risk premium, i.e. $\beta_i (E(R_m) - R_f)$. $\beta_i$ is calculated as the covariance between the returns on risky asset $i$ and the market portfolio $M$, divided by the variance of the market portfolio (see, for instance, Copeland, 2005).

\(^5\) The CAPM is also known as a single factor model since $\beta$ (i.e. the systematic risk) is the only driving risk factor.
3.3.1 Treynor Ratio

From a risk perspective there exist two different types of general CAPM decision rules. One of them suggests to choose the asset with the highest ratio of expected excess return to the systematic risk, a ratio also known as the Treynor ratio (TR) introduced by its namesake Treynor (1965). Formally, the TR is given by

\[ TR_i = \frac{E(R_i) - R_f}{\beta_i}, \]  

where \( E(R_i) - R_f \) is the expected excess return over the risk-free rate and \( \beta_i \) is the systematic risk.

Similar to the SR, the TR relates excess return to risk. However, instead of total risk, it considers only systematic risk: The higher the TR, the better is the performance under analysis.

If investors assess different investment opportunities based on the TR, they are assumed to only care about expected return and systematic risk. Therefore, a ranking of portfolios based on the TR is only useful if the investment under consideration is a sub-investment of a broader, fully diversified portfolio. If this is not the case, portfolios with identical systematic risk, but different total risk, will be rated the same. According to the CAPM, the portfolio with the higher total risk is less diversified and therefore has a higher unsystematic risk, which the market will not reward.

Hence, for the TR to be reliable, two additional conditions must hold in comparison with the SR. First, investors choose investments or portfolios according to the mean-beta framework. Second, there is only one source of risk, that alternative investments are equally correlated to.

3.3.2 Jensen’s Alpha and the Appraisal Ratio

The other type of CAPM decision rules are known as the Jensen’s alpha introduced by Jensen (1968) and the Appraisal ratio (AR) introduced by Treynor and Black (1973). Jensen’s alpha suggests to choose the investment that maximizes the abnormal return \( \alpha_i \), irrespectively to the incurred risk, and is defined as

\[ \alpha_i = R_i - [R_f + \beta_i(R_m - R_f)] , \]

where \( R_i \) is the return of asset \( i \), \( R_f \) is the risk-free rate, \( R_m \) the market return and \( \beta_i \) the coefficient which represents the systematic risk incurred in asset \( i \).

For the Jensen’s alpha rule to be a reliable RAPM, three major conditions must hold compared to the SR. The first two rules are identical to the ones of the TR (i.e. the mean-beta framework condition as well as the correlation condition). The third additional condition requires that the returns on the two investments have the same risk in order to receive an unbiased solution. Thus,
Jensen’s alpha is even more restrictive than the TR and is strictly speaking not even a RAPM, as it does not account for the risk taken.\(^6\)

The Appraisal ratio (AR) suggests to chose the asset with the highest ratio of \(\alpha\) to \(\beta\) and is defined as

\[
AR_i = \frac{\alpha_i}{\beta_i},
\]

where \(\alpha_i\) is the part of the investment’s excess return that is not explained by the market excess return and \(\beta_i\) is the systematic risk, i.e. the sensitivity of the asset’s return to changes in the market portfolio’s excess return.\(^7\)

The result is a ratio that measures the abnormal return per unit of systematic risk. In a mean-beta framework this is straightforward, as investors are only rewarded for systematic risk, whereas the conditions for delivering the correct answers are identical to the ones of the TR.\(^8\)

### 3.3.3 CAPM Performance Measures as traditional RAPMs

The traditional RAPMs include all RAPMs that are based on the mean-variance framework, implying that the investor’s preferences can be represented by an exponential utility function and that returns are normally distributed. This includes also the CAPM decision rules, as they rely on the same assumptions. Yet, while the mean-variance decision rules capture systematic and non-systematic risk, the CAPM decision rules solely capture systematic risk. Consequently, the CAPM decision rules can be viewed as a subset of the mean-variance decision rules.

The choice of whether to use mean-variance decision rules or CAPM decision rules depends on the investors risk preferences. In concrete terms, if investors are concerned about total risk, i.e. the neutral position is the risk free rate, than the mean-variance decision rules are more appropriate. In contrast, the CAPM decision rules are more appropriate if investor’s are only concerned about systematic risk, i.e. the neutral position is a benchmark portfolio.

Dowd (2001) stresses that in practice the mean-variance decision rules, and in particular the SR always provide accurate rankings when assessing alternative investment opportunities, as long as returns are normally distributed. CAPM decision rules, in contrast, only provide correct rankings in

---

\(^6\) Today “alpha opportunities” are usually measured by multi-factor return models, instead of the single factor model (the CAPM). However, it is extremely complicated to obtain a consistent ranking of investment alphas, since they are so dependent on the applied models. Readers interested in additional details may consult for example Alexander and Dimitriu (2005).

\(^7\) It should be noted that if the CAPM equilibrium holds, the abnormal return \(\alpha_i\) would always be zero and as a result the AR and the Jensen’s alpha would not provide a solution, i.e. would be zero for all assets.

\(^8\) Along with the evolution of multi-factor models, such as Fama and French (1992, 1993) a generalization of the TR has been proposed, allowing for an inclusion of the investment sensitivities to more than one factor. See e.g. Hübner (2005) for more information.
special cases or coincidently. Table 3.2 provides an illustration of how misleading the CAPM decision rules might become, when some of the underlying restrictions are violated. However, when returns are non-normally distributed, even the SR can lead to misleading conclusions and unsatisfactory paradoxes (see for instance Hodges, 1998; Bernardo & Ledoit, 2000).

<table>
<thead>
<tr>
<th>Investment</th>
<th>Market portfolio</th>
<th>Risk-free asset</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected return</td>
<td>12.0%</td>
<td>8.0%</td>
<td>9.0%</td>
</tr>
<tr>
<td>Standard deviation of return</td>
<td>12.0%</td>
<td>8.0%</td>
<td>9.0%</td>
</tr>
<tr>
<td>Correlation with the market return</td>
<td>1.0</td>
<td>0.5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 3.2: This is an illustration of the unreliability of the CAPM decision rules if the returns of alternative investment opportunities are not equally correlated with the market portfolio. According to the SR investment A should (correctly) be chosen. In contrast, both the TR and the AR suggest to choose investment B over A, while Jensen’s Alpha rather by luck suggests to chose investment A over B.

### 3.4 Downside Risk Performance Measures

Recently, there has been a growing literature on RAPMs that attempt to take into account higher moments of the return distribution. This development is driven by two major arguments: First, unlike condition (A3) investor’s perception of risk goes beyond the variance and hence investors are also averse to negative skewness and positive excess kurtosis. Second, unlike condition (A5), many investment returns have non-normally distributions. As already mentioned in chapter 2, it was Markowitz (1959) himself who noted that using the variance to measure risk is too conservative, since it regards all extreme returns – positive and negative ones – as undesirable. This claim was later validated by Lhabitant (2001) who examined the problems of RAPMs for non-normally distributed returns and concluded that the mean-variance RAPMs are inadequate for non-normal return distributions.

Motivated by the common interpretation of the SR as a reward-to-risk ratio, academics and practitioners tried to mitigate the shortcomings of the variance as a risk measure. As a result, they replaced the standard deviation in the denominator of equation (3.1) with alternative risk measures – mostly downside risk measures – in order to account for the non-normality of returns. This development was driven by the requirement to evaluate investments with odd-shaped return distributions. In particular, hedge funds are well known to be prone to generate returns that significantly deviate from normality (see, for instance, Brooks & Kat, 2002; Agrawal & Naik, 2004; Malkiel & Saha, 2005).
3.4.1 The Modified Sharpe Ratio

Motivated by the existence of non-normally distributed investment returns, Gregoriou and Gueyie (2003) proposed their Modified Sharpe ratio (MSR) as an improvement to the original SR by replacing the standard deviation in the denominator of the SR with the Modified Value-at-Risk (MVaR). The MVaR, which is based on a Cornish-Fisher expansion, is similar to the classical Value-at-Risk (VaR) but also takes into account skewness (i.e. return asymmetries) and kurtosis (i.e. fat tails) at a given confidence level $\alpha$. Hence, the MVaR is defined as

$$
MVaR = \mu - \left( z_c + \frac{1}{6} (z_c^2 - 1) \tau + \frac{1}{24} (z_c^3 - 3z_c) \kappa - \frac{1}{36} (2z_c^3 - 5z_c) \tau^2 \right) \sigma,
$$

where $\mu$ is the mean of the investment, $\sigma$ is the investment’s standard deviation, $\tau$ is the assets skewness, $\kappa$ is the investments kurtosis and $z_c$ is the critical value of the normal standard distribution at a $(1 - \alpha)$ threshold.

Since the MVaR replaces the standard deviation the MSR is defined as

$$
MSR = \frac{E(R) - R_f}{MVaR},
$$

where $E(R) - R_f$ is the excess return of an investment and $MVaR$ is equal to equation (3.9).
Academics and practitioners also proposed to use other similar RAPMs such as excess return over value at risk (see Dowd, 2001) or the conditional SR (see Agarwal & Naik, 2004). However, using the VaR or the MVaR as a risk measure for assessing different investment opportunities is controversial, since the VaR has no direct link to the utility of an investor. Artzner et al. (1997) stated that the VaR measure even violates the most basic requirements for risk measures. In fact, it is not a coherent risk measure, which requires the risk measure to weigh all quantiles above the VaR at least non-decreasing (see sections 4.4.1 and 4.4.2).

The VaR assigns, for instance, the 95% quantile a weight of 1, whereas all other quantiles get a weight of 0. This implies that investors do not care about losses that exceed a certain threshold. However, in reality this is not plausible, because higher losses should also attribute a higher weights due to the fact that investors are generally assumed to be risk averse (see Artzner et al., 1999). Even the expected shortfall (ES) represents the rather unrealistic assumption that investors are assumed to be risk neutral above the VaR.

However, the VaR was introduced rather for regulatory and risk management purposes and not for assessing investment opportunities. Be that as it may, it is still the most prominent downside risk measure, mainly due to the Basel committee on Banking Supervision who proposed the standards for minimum capital requirements for banks based on the VaR. See also chapter 4, which describes this issue in more detail.

### 3.4.2 The Sortino Ratio

In the early 1990s, Sortino and Van der Meer (1991) tried to resolve the shortcomings of the SR and therefore introduced a new RAPM, which came to be known as the Sortino ratio (SoR). The SoR is equivalent to the SR except that the square root of the lower partial moment (LPM) of second order replaces the standard deviation. It is for this reason a more realistic measure of risk-adjusted return than the original Sharpe ratio, since investors are indeed not averse against high returns, as implied by the variance. The SoR is therefore defined as

$$SoR = \frac{E(R) - \tau}{\sqrt{\int_{-\infty}^{\tau} (R - \tau)^2 dF(R)}}$$

(3.11)

where $E(R) - \tau$ is the excess return over some threshold, which is usually also referred to as the minimum acceptable return (MAR) and $F(.)$ is the cumulative density function (cdf) for total returns on an investment defined on the interval $(-\infty, \tau)$.

The LPM of second order is an asymmetric risk measure that calculates the probability-weighted squared deviations of those returns falling below a specified threshold return (i.e. to the left of the distribution). In most literature it is also referred to as semi-standard deviation and its square to the semi-variance, although such a terminology is only correct when $\tau = E(R)$. However, unlike the
variance the semi-variance is sensitive to skewness of the return distribution and the probability of shortfalls.

Accordingly, the SoR is highly vulnerable to biased data sets if the ex-post estimation is based on a period of upwardly trending returns, because the downside deviation underestimates the two-sided risk if the estimation period is not long enough to include loss periods. In this case, the SR would perform better, since the standard deviation is not as vulnerable to a skewed sample when the underlying population is symmetrical. Additional information about the general properties of this kind of RAPMs are provided in section 3.4.4.

3.4.3 Omega and Sharpe-Omega

Keating and Shadwick (2002) introduced Omega, which incorporates all the characteristics of a return distribution and does not need any assumptions about risk preferences or utilities. It was developed to overcome the inadequacy of the traditional performance measures when applied to investments with non-normal return distributions. More precisely, it considers the returns below and above a specific loss threshold and provides a ratio of total probability in weighted losses and profits that fully describe the risk-reward properties of a distribution. As a result, it can also be applied as a performance measure. Keating and Shadwick (2002) even claimed that the Omega function is a universal performance measure as it represents a complete description of the return distribution. Formally, the Omega function is defined as

\[
\Omega(\tau) = \frac{\int_{\tau}^{\infty}[1 - F(R)]dR}{\int_{-\infty}^{\tau}F(R)dR},
\]

where \( \tau \) is a threshold return selected by an investor and \( F(\cdot) \) is the cumulative density function (cdf) for total returns on an investment. For any investor, returns below a specific threshold are considered as losses and returns above as gains. At a given threshold, a higher value of Omega is preferred to a lower one.

From a statistical point of view, Omega is equivalent to the return distribution itself and embodies all its moments. As the Omega function is a unique, monotone, decreasing function of the cumulative distribution of returns, its first order derivative is always negative. Hence, the Omega function of a risky distribution is flatter than that of a less risky distribution. The mean return of the distribution represents the unique point at which the Omega function takes the value of 1, which is the point where the total weighted probability of gains are equal to the one of the losses.
Shortly after the Omega was introduced, Kazemi, Schneeweis and Gupta (2004) showed that it is not significantly new in finance, as it represents a ratio of a call option to a put option with the strikes at the same threshold, i.e.\(^9\)

$$\Omega(\tau) = \frac{C(\tau)}{P(\tau)} = \frac{E[\max(R - \tau, 0)]}{E[\max(\tau - R, 0)]}, \quad (3.13)$$

where \(C(\tau)\) is the price of an European call option written on an investment and \(P(\tau)\) the price of a European put option written on the same investment\(^{10}\). The strike price of both options is equal to the threshold return \(\tau\) and their maturity is equal to one period (e.g., 1 year).

**Figure 3.3:** This is a graphical representation of Omega, where the red line represents the cumulative density function of the returns, and \(\tau\) is the given threshold. Omega is calculated as the quotient of the density above the threshold over the density below the threshold.

Kazemi et al. (2004) also introduced a new RAPM called the Sharpe-Omega. Sharpe-Omega is defined as

$$\text{Sharpe}_\Omega = \frac{E(R) - \tau}{P(\tau)} = \frac{E(R) - \tau}{E[\max(\tau - R, 0)]}, \quad (3.14)$$

where \(E(R) - \tau\) is the expected excess return on the investment given threshold \(\tau\), and \(P(\tau)\) is the price of the put option on the same investment given strike \(\tau\). Moreover, Kazemi et al. (2004) showed that equation (3.14) is equal to

\(^9\) \(E[\max(R - \tau, 0)]\) and \(E[\max(\tau - R, 0)]\) are undiscounted call and put prices. In line with the literature the concept is, for convenience reasons, presented without considering the discount factors. For an exact calculation see Kazemi et al. (2004).

\(^{10}\) The interested reader is referred to Kazemi et al. (2004), which provide in their appendix a detailed proof for this relation.
and therefore provides the same information as the Omega measure (see figure 3.4).

The Sharpe-Omega is a measure that is more intuitive than Omega, since it is also some sort of risk-reward ratio, where the price of the put option captures the riskiness of an investment as the costs of protecting the investment from returns below the threshold. Thus Sharpe-Omega ranks investments according to their return divided by the costs to protect this return.

If, for example, an investment has a threshold higher than the expected return, the Sharpe-Omega would be negative. In that case, a higher put price, caused for instance by a higher volatility, would be, according to the Sharpe-Omega result a better investment opportunity. The opposite holds if the return threshold is below the expected return.

Figure 3.4: The black curve is a plot of the Omega ratio for a standard normally distributed \( \sim N(0,1) \) variable, whereas the red curve represents the Shape-Omega ratio for \( \sim N(0,1) \) variable. The graphical representation illustrates that \( \Omega(\tau) - 1 \) is indeed equal to the Sharpe-Omega ratio.

In conclusion, Omega and Sharpe-Omega offer flexible and easy to use alternative RAPMs. Their non-parametric nature reduces the potential for errors and makes the ratios as statistically significant as the data they are assessing. They can also be plotted on a log scale to differentiate between even the most extreme asymmetries. Thus, they are ideally suited for assessing the performance of hedge funds\(^{11}\).

\(^{11}\) The reader interested in how to construct Omega portfolios is referred to Alexander (2005), who provides a detailed overview of this subject.
3.4.4 The Kappa Indices

Kaplan and Knowles (2004) introduced with the Kappa indices a generalized downside RAPM that has lower partial moments (LPM) in the denominator. They also showed that both the SoR and the Omega are special cases of the Kappa indices, where a single parameter determines whether the SoR, Omega, or another RAMP is generated. Kappa of order $n$ is given by

$$K_n(\tau) = \frac{E(R) - \tau}{[LPM_n(\tau)]^{\frac{1}{n}}}.$$  \hspace{1cm} (3.16)

where $\tau$ is some threshold return, $n > 0$ usually an integer and the LPMs of order $n \geq 0$ are defined by Harlow (1991) as

$$LPM_n(\tau) = \int_{-\infty}^{\tau} (\tau - R)^n dF(R) = E[max(0, \tau - R)^n].$$  \hspace{1cm} (3.17)

When substituting equation (3.17) into equation (3.11), a fully equivalent definition of the Sortino ratio is obtained, i.e.

$$SoR = \frac{E(R) - \tau}{[LPM_2(\tau)]^{\frac{1}{2}}} = \frac{E(R) - \tau}{E[max(\tau - R, 0)^2]^{\frac{1}{2}}}. \hspace{1cm} (3.18)$$

Accordingly, the Kappa indices can be used as a generalization of this relation, which is given as

$$K_n(\tau) = \frac{E(R) - \tau}{[LPM_n(\tau)]^{\frac{1}{n}}}. \hspace{1cm} (3.19)$$

Evidently, $K_2(\tau)$ is identical to the Sortino ratio where the threshold is equal to an investor’s target return, whereas $K_1(\tau)$ is identical to the Sharpe-Omega ratio\(^{12}\), i.e.

$$\Omega(\tau) - 1 = \frac{E(R) - \tau}{P(\tau)} = \frac{E(R) - \tau}{LPM_1(\tau)} = \frac{E(R) - \tau}{E[max(\tau - R, 0)]}. \hspace{1cm} (3.20)$$

Although the SoR and Omega seem to be distinct mathematically, they are in fact conceptually related downside RAPMs. After all, the Kappa indices include Omega and the SoR, and thus have identical structures. They all increase as the threshold decreases. They are negative when $\tau > E(R)$, zero when $\tau = E(R)$ and positive when $\tau < E(R)$. When returns are normally distributed all Kappa indices give the same rankings compared to the SR, whereas a higher value of Kappa is always better than a lower value.

The only difference is that higher order Kappa indices deviate from lower order ones, as they are more sensitive to skewness and excess kurtosis, due to their higher sensitivity to extreme returns\(^{13}\).

\(^{12}\) This is in line with the findings of Kazemi et al. (2004) shown in Section 3.4.3

\(^{13}\) The interested reader is referred to Kaplan and Knowles (2004), who provide a detailed analysis of the sensitivities of the Kappa indices to skewness and excess kurtosis.
In contrast are higher order Kappa indices less sensitive to the choice of the threshold (see figure 3.5).

![Graph](image)

**Figure 3.5:** This is a representation of Kappa 1, 2, 3 and 4 with respect to different thresholds (the thresholds appear on the x-axis and the kappa values on the y-axis). It clearly shows that there is an inverse relationship between the threshold and the value of Kappa. The steepness of the Kappa curve decreases when the parameter $n$ increases. Thus, the most sensitive Kappa index with respect to changes of the threshold is Kappa 1.

Kaplan and Knowles (2004) have shown that if shape estimation functions are used for investment return distributions, it is sufficient to estimate the first four moments of a return distribution in order to obtain robust Kappa indices. Hence, in order to calculate the Kappa indices, in a parametric set-up, the application the Johnson family of distributions (Stuart & Ord, 2004; Johnson, 1949) is ideally suited, as it is uniquely determined by the first four moments (i.e. mean, unit variance, skewness, and kurtosis).

Since the Kappa indices can be viewed somewhat as the mother of all downside RAPMs, there is currently a heated debate among academics and practitioners over the appropriate choice of the index and the threshold in order to rank different investment opportunities. This is of crucial importance as the choice of one Kappa index over another might have a significant effect on the rankings of alternative investment opportunities. However, there is no generally applicable rule for choosing the best Kappa index, as it cannot be linked to a standard utility function. Alexander (2008a) stated the utility linked to the Kappa indices is somewhat strange, because “it can be anything at all for ‘upside’ returns that are above the threshold” (p. 265) 14.

---

14 See also the paradox of the lottery problem stated by Shadwick (2004), which examines whether it is better to buy a lottery ticket for $1 or to sell it. The excess kurtosis and skewness are extreme and only Omega ranks the buy decisions better than sell for all thresholds. The SoR, in contrast, fails the test, and for all kappa indices $n>2$ it depends on the threshold whether they rank the buy decision higher than the sell one.
All that can be said about the Kappa index is that those investors who are more risk averse to skewness and excess kurtosis of downside returns should choose a high order Kappa index in order to rank their investment opportunities. On the other hand the question of how to choose the threshold remains unanswered. Known is the following: As pointed out by Koekebakker and Zakamouline (2009) and also concluded by Ornelas et al. (2009), rankings based on Kappa indices are highly sensitive to the choice of this threshold. While some argue that the risk-free rate should be chosen, others suggest the threshold to be very high, since investors are assumed to be indifferent to returns above the threshold.

### 3.4.5 The Upside Potential Ratio

Only a few years after Sortino and Van der Meer (1991) introduced the SoR, Sortino, van der Meer and Plantinga (1999) suggested another RAPM called the upside potential ratio (UPR), which is somewhat similar to Omega but uses instead of the LPM of first order the square root of the LPM of second order (i.e. semi-standard deviation, where $\tau = E[R]$).

The UPR measures the extent to which an investor has been able to achieve upside potential relative to downside risk, with minimum acceptable return (MAR) being the reference point. Hence, the UPR is often also called upside potential to downside risk ratio. Sortino et al. (1999) defined the UPR as

$$\text{UPR}(\tau) = \frac{\int_\tau^\infty (R - \tau) f(R) dr}{\left[ \int_\tau^\infty (R - \tau)^2 f(R) dr \right]^{\frac{3}{2}}}$$

mentioned before, the UPR may also be expressed as a ratio of partial moments

$$\text{UPR}(\tau) = \frac{\text{HPM}_1(\tau)}{[\text{LPM}_2(\tau)]^{\frac{3}{2}}}$$

where $\tau$ stands for the MAR of an investor or a fund manager, and $\text{HPM}_1$ is the higher partial moment (HPM) of the first order.

The UPR was mainly introduced to judge a fund manager’s ability to accomplish the goals of investors, which Sortino et al. (1999) argue is best achieved by measuring the performance relative to the MAR. They based their claim on the emerging field of Behavioral Finance and the esoteric area of the utility theory. They cited, for instance, Tversky (1995), who proposed the prospect theory reflecting the behavior of investors more adequately. Essentially, the assumption that investors are rational is disputed, because they are usually not; while they, for instance, are normally risk averse in gains, they are usually risk-seeking in losses, which is hardly rational.
The bottom-line is that most investors seek wealth growth that is as stable as possible, hence Sortino et al. (1999) claim that most investors prefer fund managers who achieve the highest average return over their MAR, and not the ones with the highest return over some period. The UPR measures the fund manager’s style by combining the potential for success with the risk of failure (i.e., the upside potential to the downside risk) and enables choosing strategies with growth that is as stable as possible for a given minimum return.

### 3.4.6 RAPMs based on Maximum Drawdown

Drawdown based RAPMs are fairly popular in practice particularly in the world of commodity trading advisors and hedge funds, owed to the widespread use of the Calmar ratio proposed by Young (1991), the Sterling ratio proposed by Kestner (1996) and the Burke ratio introduced by Burke (1994). They are all computed similarly as the annual average excess return divided by a variant of the maximum drawdown (i.e., the maximum loss incurred over a given time period, usually quoted as the percentage between the peak and the low). Consequently, the higher either of these ratios, the better the portfolio performance. Hence, it is another approach that tries to capture odd-shaped return distributions.

Readers interested in a precise calculation are referred to the cited literature or, for instance, to Eling and Schuhmacher (2007).

### 3.5 Preference-Based Performance Measures

Downside RAPMs have frequently been criticized, as their application has three major shortcomings: First, their ranking depends heavily on the chosen threshold. Second, they only take abnormal downside risk into account, whereas potential abnormal upside returns (i.e., positive skewness) are usually neglected. And third, while the mean-variance RAPMs are based on the expected utility theory - the cornerstone of modern finance - these RAPMs have often no plausible relation with risk-averse utility functions. Therefore, downside RAPMs are likely to be misleading (see, for instance, Pézier, 2008).

As a result, there is a growing amount of literature in recent performance measurement journals that derives the RAPM from utility functions characterizing investor’s preference for higher moments. This new class of RAPMs is especially useful to assess the attractiveness of investments with non-normally distributed returns. Unfortunately, they often require numerical methods in order to calculate them and are, moreover, not as easy to understand as most other RAPMs. This section introduces the most popular RAPMs that explicitly account for the investor’s risk preference through the use of a utility function.
3.5.1 The Generalized Sharpe Ratio

Hodges (1998) introduced a generalization of the SR that takes into account all moments of a distribution and, moreover, is linked to the investor’s utility function. This RAPM came to be known as Generalized Sharpe ratio (GSR). The GSR has the advantage that it can accommodate many distributions with non-normal features such as discontinuities, multiple peaks, major skewness, etc. In order to handle such distributions, however, the GSR needs to make assumptions about investor’s behavior. More precisely, the investor’s utility function is assumed to be exponential, and investors are assumed to always be able to find the maximum utility of a portfolio. Given these assumptions, where \( E^* \) is the maximum of the expected utility, the GSR of a portfolio is defined as

\[
GSR = (-2\ln(-E^*))^{\frac{1}{2}}. 
\]  
(3.23)

While the GSR has convenient properties, it is usually more complex to calculate than other RAPMs, as it requires the maximum expected utility to be computed. However, there is an exception when investors have exponential utility functions. In that case, the maximum expected utility can – according to Pézier (2008) – be approximated by the fourth Taylor approximation of the certain equivalent excess return (CER), i.e.

\[
CER(q) \approx q\mu - \frac{1}{2} \gamma \sigma^2 q^2 + \frac{1}{6} \tau \gamma^2 \sigma^3 q^3 - \frac{1}{24} \kappa \gamma^3 \sigma^4 q^4, 
\]  
(3.24)

where the mean \( \mu \), standard deviation \( \sigma \), skewness \( \tau \) and excess kurtosis \( \kappa \) correspond to the excess returns when a unit amount is invested in the portfolio, \( \gamma \) expresses the coefficient of risk aversion (as proportion of the amount invested) and \( q > 1 \) represents the multiplicative factor, which increases the invested amount in the portfolio. The maximum CER in equation (3.24) can be found by a numerical optimization, which in turn can be set equal to the maximum expected utility in equation (3.23) in order to find the GSR.

Instead of numerical optimization, Pézier (2008) uses another approach to calculate the maximum CER by approximating the multiplicative factor \( q \) as

\[
q^* \approx \frac{\mu}{\gamma \sigma^2}, 
\]  
(3.25)

He shows that if \( q \) is equal to equation (3.25), and given that the investor has an exponential utility function, the maximum expected utility is equal to

\[
E^* \approx -\exp \left( -\frac{1}{2} \left( \frac{\mu}{\sigma} \right)^2 - \frac{1}{6} \tau \left( \frac{\mu}{\sigma} \right)^3 + \frac{1}{24} \kappa \left( \frac{\mu}{\sigma} \right)^4 \right). 
\]  
(3.26)

Because \( \mu/\sigma = \lambda \) is equal to the ordinary SR where \( \mu \) is the excess return, the equation can be rearranged as
\[ -2 \ln(-EU^*) \approx \lambda^2 + \frac{1}{3} \tau \lambda^3 - \frac{1}{12} \kappa \lambda^4. \]  \hspace{1cm} (3.27)

Hence, the GSR given an exponential utility of an investor is equal to the square root of equation (3.27), i.e.

\[ GSR \approx \left( \lambda^2 + \frac{1}{3} \tau \lambda^3 - \frac{1}{12} \kappa \lambda^4 \right)^{\frac{1}{2}}. \]  \hspace{1cm} (3.28)

Assuming that the utility function of an investor is exponential and the returns are normally distributed, the GSR is identical to the ordinary SR, i.e.

\[ GSR \approx (\lambda^2)^{\frac{1}{2}}. \]  \hspace{1cm} (3.29)

If returns are non-normally distributed, say the distribution is left skewed and has a positive excess kurtosis, this will result in a reduction of the GSR, while the ordinary SR will be unaffected.

Pézier (2008) states that basically all traditional RAPMs such as the SR, IR, and Jensen’s alpha can be generalized by maximizing the CER. This approach is, moreover, consistent with the expected utility theory and thus applicable to any assessment of investment opportunities. The CER is quoted in basis points, so it is easy to interpret and does not impose any restriction on the return distribution and also takes into account the investor’s risk attitude. It even reveals the composition of the corresponding optimal investment portfolio.

### 3.5.2 Alternative Investments Risk-adjusted Performance

Sharma (2004) introduced with the alternative investments risk-adjusted performance (AIRAP) measure a RAPM dedicated to the assessment of hedge funds. He uses an approach similar to the one of Pézier (2008) by calculating a risk-adjusted certainty equivalent (CE) under the assumption of a constant relative risk aversion (CRRA), representing the investor’s risk preferences. In Sharma’s words the CE “is the implied equivalent return that the risk-averse investor desires with certainty in exchange for holding uncertain risky assets” (p. 3). The price paid for trading off the risky investment with its CE (i.e. the AIRAP) is equal to the risk premium \( \text{RP}(z) = [E(z) - CE(z)] \), where \( \text{RP}(z) \geq 0, z \) are the payoffs. Figure 3.6 shows the standard RP > 0 for risk averse investors, which corresponds to a CRRA > 0 and a concave utility function, respectively\(^{15}\).

Sharma (2004) recommends the use of a CRRA on the interval \([2, 4]\), even though a range from 1 to 10 is considered as possible.

\(^{15}\text{CRRA} = 0 \) corresponds to a risk neutral investor with a linear utility function, while \( \text{CRRA} < 0 \) corresponds to a risk loving investor with a convex utility function.
Formally, Sharma (2004) defines the AIRAP as follows:

\[
\text{AIRAP} = \left[ \sum_i p_i (1 + TR_i)^{(1-c)} \right]^{1/(1-c)} - 1, \quad \text{when } c \neq 1 \text{ & } \geq 0
\]  

(3.30)

and \( \text{AIRAP} = \prod_i (1 + TR_i)^{p_i} - 1, \) \( \text{when } c = 1 \)  

(3.31)

where \( p_i \) is the probability of the return \( i \), \( TR_i \) is the period total fund return \( i \), \( c \) = CRRA parameter, \( i = \text{index } 1,...,N \) and \( N = \text{number of periods} \).

Sharma (2004) names several advantages of the AIRAP. First, it captures all observed higher moments and works even when mean returns are negative. Second, it punishes high volatility and high leverage in proportion to the risk aversion of the investor. Third, it is in its closed form solution as simple to calculate as the original SR, while limitations of the mean-variance framework concerning non-normal returns are avoided.

![Figure 3.6](image)

Figure 3.6: This represents the idea of AIRAP (CE) under CRRA. The blue curve represents a power utility function for CRRA > 0, i.e. a concave utility function, implying risk aversion. \( E(z) \) is a combination of the returns \( Z_1 \) and \( Z_2 \) and the certainty equivalent corresponds to the AIRAP. Hence, the risk averse investor desires a risk premium in the form of \( E(z) \)-AIRAP in exchange for holding the risky investment.

### 3.5.3 The Manipulation Proof Performance Measure

Essentially all RAPMs introduced so far have been designed to capture the performance of actively managed funds. Yet if a specific RAPM is being used to assess the performance of a portfolio manager, then the manager most certainly has an incentive to manipulate the applied measure, as
his income is usually based on this measure. Unfortunately, such manipulation activities (sometimes also referred to as gaming activities) appear to have substantial impact on some of the most popular RAPMs (see section 3.7). Hence, Goetzmann et al. (2007) investigated the question of whether a RAPM exists that cannot be manipulated. As a result, Goetzmann et al. (2007) proposed four conditions that a RAPM must satisfy to become a manipulation proof performance measure:

(B1) the performance measure must be increasing with the return,
(B2) the function must be concave,
(B3) it must be time separable,
(B4) and it must have a power utility form.

The first condition relates to the property that a RAPM must recognize arbitrage opportunities. This implies that the measure must recognize if a return distribution \( r_1 \) dominates another return distribution \( r_2 \) and thus must rank it higher, i.e. \( \Theta(r_1) > \Theta(r_2) \), where \( r \) is a vector of returns across all the possible outcomes and \( \Theta(r) \) is the performance measure as a function of those returns. The second condition prevents the measure from increasing by simply changing the leverage or adding potential but unpriced risk, hence condition (B1) and (B2) prevent static manipulations, which is also consistent with stochastic dominance of first order. The third condition requires the performance measure to be free from dynamic manipulations, which is satisfied by simple performance measures such as the geometric average return.

Any measure that satisfies conditions (B1) to (B3) is manipulation proof but not necessarily consistent with the market equilibrium. Hence, the fourth condition requires the measure to be consistent with the economic equilibrium, which is only true for the power utility function. Taking into account all these conditions, Goetzmann et al. (2007) derived their own manipulation proof performance measure (MPPM), i.e.

\[
\text{MPPM} \equiv \tilde{\Theta}(p) \equiv \frac{1}{(1-p)\Delta t} \ln \left( \frac{1}{T} \sum_{t=1}^{T} \left[ \frac{(1 + r_t)/(1 + r_{ft})}{1 - e^{T \Delta t}} \right]^{1-p} \right),
\]

(3.32)

where \( p > 1 \) corresponds to the CRRA (see appendix for proof), \( \Delta t \) is the length of time between the observations (e.g. \([1/12]\) for monthly returns), \( T \) is equal to the number of total observations, \( r_t \) is the realized return at time \( t \) (not annualized) and \( r_{ft} \) is the risk-free rate at time \( t \) (not annualized).

In general the MPPM can be interpreted as the annualized continuously compounded excess CER of the portfolio. Hence, if a risk free portfolio earns \( \exp[\ln(1 + r_{ft}) + \tilde{\Theta} \Delta t] \) each period, it would have the same performance as the portfolio \( \tilde{\Theta} \).
The MPPM is not based on a specific underlying distribution and is also fairly simple to compute. In addition it penalizes negative excess returns more as the CRRA increases. Similar to the AIRAP, it also suggests to choose a CRRA of \( p \) on the interval \([2, 4]\).

Table 3.3 and table 3.4 compare the rankings of the SR, the GSR and the MPPM with 5 simulated scenarios. While in table 3.3 the MPPM is calculated with a CRRA of 4, in table 3.4 it is calculated with a CRRA of 10, implying a highly risk averse investor. The tables show that the results significantly vary depending on the level of the CRRA. For instance, scenario 1 is ranked on the 5\(^{th}\) position applying a CRRA of 4, whereas it is ranked on 1\(^{st}\) position when applying a CRRA of 10. This is consistent with the utility function of an average and a highly risk-averse investor, because high standard deviation, negative skewness and positive excess kurtosis are being stressed more, as the CRRA increases. This is illustrated in figure 3.7, which shows how the power utility function alters when a CRRA of 10 is applied instead of a CRRA of 4. Moreover, this comparison also reveals the vulnerabilities of the SR, if returns are not normally distributed and in particular if the distribution has a high positive skewness.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>X-Return</th>
<th>Std</th>
<th>Skew</th>
<th>X-Kurt</th>
<th>Min</th>
<th>Max</th>
<th>SR</th>
<th>Rank</th>
<th>GSR Rank</th>
<th>MPPM Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.83%</td>
<td>8.37%</td>
<td>2.0470</td>
<td>-1.82%</td>
<td>33.09%</td>
<td>0.3384</td>
<td>5</td>
<td>0.3706</td>
<td>3</td>
<td>1.75%</td>
</tr>
<tr>
<td>2</td>
<td>1.98%</td>
<td>5.54%</td>
<td>1.5191</td>
<td>-1.08%</td>
<td>14.40%</td>
<td>0.3567</td>
<td>4</td>
<td>0.3867</td>
<td>1</td>
<td>1.44%</td>
</tr>
<tr>
<td>3</td>
<td>2.78%</td>
<td>7.21%</td>
<td>-1.4215</td>
<td>0.3777</td>
<td>-13.87%</td>
<td>7.13%</td>
<td>0.3852</td>
<td>2</td>
<td>0.3473</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>2.56%</td>
<td>6.49%</td>
<td>-1.2214</td>
<td>4.2905</td>
<td>-46.49%</td>
<td>19.23%</td>
<td>0.3947</td>
<td>1</td>
<td>0.3494</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>2.55%</td>
<td>6.89%</td>
<td>0.1489</td>
<td>0.2128</td>
<td>-18.90%</td>
<td>26.93%</td>
<td>0.3708</td>
<td>3</td>
<td>0.3737</td>
<td>2</td>
</tr>
</tbody>
</table>

**Table 3.3:** This is an illustration of the MPPM under a CRRA of 4. The data shows 5 simulated scenarios of return distributions with distinct characteristics (i.e. excess return, standard deviation, skewness and excess kurtosis). It is obvious that the applied RAPMs provide different rankings, even though the GSR and also MPPM take into account higher moments. However, both the GSR and the MPPM rank scenario four at 4\(^{th}\) ranks, while the SR rank it as 1\(^{st}\), as the neglects the negative skewness and positive excess kurtosis.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>X-Return</th>
<th>Std</th>
<th>Skew</th>
<th>X-Kurt</th>
<th>Min</th>
<th>Max</th>
<th>SR</th>
<th>Rank</th>
<th>GSR Rank</th>
<th>MPPM Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.83%</td>
<td>8.37%</td>
<td>2.0470</td>
<td>-1.82%</td>
<td>33.09%</td>
<td>0.3384</td>
<td>5</td>
<td>0.3706</td>
<td>3</td>
<td>0.78%</td>
</tr>
<tr>
<td>2</td>
<td>1.98%</td>
<td>5.54%</td>
<td>1.5191</td>
<td>-1.08%</td>
<td>14.40%</td>
<td>0.3567</td>
<td>4</td>
<td>0.3867</td>
<td>1</td>
<td>0.85%</td>
</tr>
<tr>
<td>3</td>
<td>2.78%</td>
<td>7.21%</td>
<td>-1.4215</td>
<td>0.3777</td>
<td>-13.87%</td>
<td>7.13%</td>
<td>0.3852</td>
<td>2</td>
<td>0.3473</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>2.56%</td>
<td>6.49%</td>
<td>-1.2214</td>
<td>4.2905</td>
<td>-46.49%</td>
<td>19.23%</td>
<td>0.3947</td>
<td>1</td>
<td>0.3494</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>2.55%</td>
<td>6.89%</td>
<td>0.1489</td>
<td>0.2128</td>
<td>-18.90%</td>
<td>26.93%</td>
<td>0.3708</td>
<td>3</td>
<td>0.3737</td>
<td>2</td>
</tr>
</tbody>
</table>

**Table 3.4:** This is an illustration of the MPPM under a CRRA of 10. The data exhibits how the CRRA of a highly risk averse investor affects the rankings of the scenarios. High standard deviation, negative skewness and positive excess kurtosis are highly penalized. Hence, such an investor would prefer the scenario with the least risk, but also the lowest expected excess return.
Figure 3.7: This figure compares two power utility functions. The graph on the left side shows a CRRA of 4 that is fairly common for a rather risk-averse investor. The graph on the right side illustrates an even more risk-averse investor with a CRRA of 10. This is clearly recognizable due to the high risk premium the investor requires in exchange for holding a risky investment.

### 3.6 Manipulation of RAPMs

A number of recent papers have addressed the issue of whether particular RAPMs are prone to manipulation attempts, with the result that some major RAPMs are in fact vulnerable to manipulation activities (see Leland, 1999; Spurgin, 2001; Goetzmann et al., 2002). Despite the development of numerous RAPMs that were designed redeem the SR, the SR is still the most renowned and also most commonly used measure of financial performance. Even in the hedge fund industry, where returns are well known to be non-normally distributed, is the SR, according to Goetzmann et al. (2007), doing on average a fairly good job.

Consequently, academics usually examine whether the SR is contingent to manipulations. Among those academics were Leland (1999) and Spurgin (2001), who show that portfolio managers are indeed able to increase the SR, by simply selling off the upper tail of the return distribution at the expense of a fat left tail (see table 3.5 as well as figure 3.8).

Goetzmann et al. (2007) were even able to verifying two basic manipulation strategies that enable a manager to enhance the applied RAPM, without requiring skill or additional information. The first strategy is referred to as the so-called static manipulation strategy where portfolio managers try to manipulate the shape of the underlying distribution in order to enhance the applied RAPM. This strategy involves selling options on the underlying assets in order to change the shape of the return distribution in a manner that maximizes the SR. The second strategy is based on the so-called dynamic manipulation strategy that induces temporal effects and estimation errors in order to increase the applied RAPM. In particular the assumption that the reported returns are i.i.d. can be violated.

Moreover, Goetzmann et al. (2007) illustrated in their study that most traditional RAPMs (i.e. SR, TR, AP, $M^2$ and the IR) are subject to the same type of manipulation that is being used on the SR.
This stem from the fact, that all traditional RAPMs are based on the assumptions of the mean-variance framework.

Goetzmann et al. (2007) also investigated whether RAPMs such as the SoR and the UPR are prone to manipulations. Even though they avoid penalizing very good returns such as the SR, they do this even too well and are therefore also subject to manipulations. A manager might be attracted to minimize the downside risk, while trying to have a small probability of very high returns, which may result in an infinite ratio (this is the case if the downside risk is zero).

Table 3.5: This table compares two uncertain investments A and B. They are identical except that investment A has 1% chance for excess return of 35%, whereas investment B has 1% chance for excess return of 45%. According to the SR an investor should prefer investment A to investment B. Accordingly, any rational investor should choose investment B over investment A. However, since the variance is a symmetric risk measure, also positive deviations from the mean are included in the risk measure (i.e. the standard deviation). This example shows that the SR does not even respect first-order stochastic dominance (see Levy & Robinson, 1998). In fact, all RAPMs using a symmetric risk measure are subject to this issue.

<table>
<thead>
<tr>
<th>Probability</th>
<th>0.01</th>
<th>0.04</th>
<th>0.25</th>
<th>0.40</th>
<th>0.25</th>
<th>0.04</th>
<th>0.01</th>
<th>Sharpe Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess return investment A</td>
<td>-0.25</td>
<td>-0.15</td>
<td>-0.05</td>
<td>0.05</td>
<td>0.15</td>
<td>0.25</td>
<td>0.35</td>
<td>0.5000</td>
</tr>
<tr>
<td>Excess return investment B</td>
<td>-0.25</td>
<td>-0.15</td>
<td>-0.05</td>
<td>0.05</td>
<td>0.15</td>
<td>0.25</td>
<td>0.45</td>
<td>0.4931</td>
</tr>
</tbody>
</table>

Source: Hodges (1998)

Figure 3.8: This is a graphical representation of the example in Table 3.5. The figure on the left shows the expected return distribution of investment A, while the figure on the right shows the expected return distribution of investment B. The distribution of investment B exhibits a higher positive skewness than investment A, clearly recognizable from the right tail.

Solely RAPMs derived from a utility function have shown not to be subject to static manipulations. More precisely, Zakamouline (2009) has shown that the GSR can mitigate the paradoxes of the traditional SR and even reveals the true performance of portfolios with manipulated SRs. Similarly, Sharma (2004) found that the AIRAP is also unaffected by static manipulations, moreover
Goetzmann et al. (2007) and Brown et al. (2010) have shown that the MPPM is even unaffected by any manipulation strategy.

## 3.7 Summary of RAPMs

This section provides an overview of all introduced RAPMs in this chapter and points out the main advantages and disadvantages of the four introduced categories, i.e. the mean-variance, the CAPM, the downside risk and the preference-based performance measures. It also provides the reader with the formulas for the calculation of each introduced RAPM as well as the name of the corresponding author(s).

<table>
<thead>
<tr>
<th>Category</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean-Variance Performance Measures</td>
<td>• Easy to calculate.</td>
<td>• Inaccurate in particular when return distributions exhibit high positive skewness.</td>
</tr>
<tr>
<td></td>
<td>• Widespread acceptance.</td>
<td>• Positive deviations are punished.</td>
</tr>
<tr>
<td></td>
<td>• Have shown to provide adequate rankings, even when non-normally distributed returns are assessed.</td>
<td>• Prone to manipulations.</td>
</tr>
<tr>
<td>CAPM Performance Measures</td>
<td>• Easy to calculate.</td>
<td>• Only reliable in special cases.</td>
</tr>
<tr>
<td></td>
<td>• Adequate if the CAPM equilibrium holds.</td>
<td>• Inaccurate if return distributions are odd-shaped.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Positive deviations are punished.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Prone to manipulations.</td>
</tr>
<tr>
<td>Downside Risk Performance Measures</td>
<td>• Take into account higher moments of a return distribution.</td>
<td>• Missing link to the utility theory.</td>
</tr>
<tr>
<td></td>
<td>• Reflect the risk-averse attitude of investors (as investors are not averse to upside deviations).</td>
<td>• Highly sensitive to the chosen threshold.</td>
</tr>
<tr>
<td></td>
<td>• Approximations possible.</td>
<td>• Usually do not take into account abnormal upside returns.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Prone to manipulations.</td>
</tr>
<tr>
<td>Preference-Based Performance Measures</td>
<td>• Reflect the risk preferences of an investor.</td>
<td>• Fairly complex.</td>
</tr>
<tr>
<td></td>
<td>• Accurate even if return distributions are odd-shaped.</td>
<td>• Requires all data points (does not include the approximation of the GSR proposed by Pézier, 2008)</td>
</tr>
<tr>
<td></td>
<td>• Resistant to most manipulation attempts.</td>
<td>• Usually numerical methods are required for the calculation.</td>
</tr>
</tbody>
</table>
### Mean-Variance Performance Measures

<table>
<thead>
<tr>
<th>RAPM</th>
<th>Formula</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharpe Ratio</td>
<td>$SR_i = \frac{E(R_i) - R_f}{\sigma_i}$</td>
<td>Sharpe (1966)</td>
</tr>
<tr>
<td>Information Ratio</td>
<td>$IR_i = \frac{R_i - R_f}{\sigma_i - r_f}$</td>
<td>Zenios (1993), Zenios and Kang (1993) and Sharpe (1994)</td>
</tr>
<tr>
<td>M-squared</td>
<td>$M_i^2 = \frac{R_i - R_f}{\sigma_i} \sigma_m + R_f$</td>
<td>Modigliani and Modigliani (1997)</td>
</tr>
</tbody>
</table>

### CAPM Performance Measures

<table>
<thead>
<tr>
<th>RAPM</th>
<th>Formula</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treynor Ratio</td>
<td>$TR_i = \frac{R_i - R_f}{\beta_i}$</td>
<td>Treynor (1966)</td>
</tr>
<tr>
<td>Jensen’s Alpha</td>
<td>$\alpha_i = R_i - [R_f + \beta_i(R_m - R_f)]$</td>
<td>Jensen (1968)</td>
</tr>
<tr>
<td>Appraisal Ratio</td>
<td>$AR_i = \frac{\alpha_i}{\beta_i}$</td>
<td>Treynor and Black (1973)</td>
</tr>
</tbody>
</table>

### Downside Risk Performance Measures

<table>
<thead>
<tr>
<th>RAPM</th>
<th>Formula</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modified Sharpe Ratio</td>
<td>$MSR = \frac{E(R) - R_f}{MVaR}$</td>
<td>Gregoriou and Gueyie (2003)</td>
</tr>
<tr>
<td>Sortino Ratio</td>
<td>$SoR = \frac{E(R) - \tau}{\sqrt{\int_{-\infty}^{\tau} (\tau - R)^2 dF(R)}}$</td>
<td>Sortino and Van der Meer (1991)</td>
</tr>
<tr>
<td>Omega</td>
<td>$\Omega(\tau) = \int_{-\infty}^{\tau} [1 - F(R)]dR / \int_{-\infty}^{\tau} F(R)dR$</td>
<td>Keating and Shadwick (2002)</td>
</tr>
<tr>
<td>Sharpe-Omega</td>
<td>$Sharpe_{\Omega} = \Omega(\tau) - 1$</td>
<td>Kazemi, Schneeweis and Gupta (2004)</td>
</tr>
<tr>
<td>Kappa Indices</td>
<td>$K_{\alpha}(\tau) = \frac{E(R) - \tau}{[LPM_{\alpha}(\tau)]^{\frac{1}{2}}}$</td>
<td>Kaplan and Knowles (2004)</td>
</tr>
<tr>
<td>Upside Potential Ratio</td>
<td>$UPR(\tau) = \frac{HPM_{1}(\tau)}{[LPM_{2}(\tau)]^{\frac{1}{2}}}$</td>
<td>Sortino, van der Meer and Plantinga (1999)</td>
</tr>
</tbody>
</table>

### Preference Based Performance Measures

<table>
<thead>
<tr>
<th>RAPM</th>
<th>Formula</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSR</td>
<td>$GSR = (-2\ln(-EU^*))^{\frac{1}{2}}$</td>
<td>Hodges (1998)</td>
</tr>
<tr>
<td>AIRAP</td>
<td>$AIRAP = \prod_{i} p_i (1 + TR_i)^{(1-c)}^{\frac{1}{(1-c)}} - 1$</td>
<td>Sharma (2004)</td>
</tr>
<tr>
<td>MPPM</td>
<td>$(p) \equiv \frac{1}{(1-p)\Delta t} \ln\left(\frac{\prod_{i=1}^{t} (1 + r_i)/(1 + r_f)}{(1 + r_f/(1 + r_f))^{1-p}}\right)$</td>
<td>Goetzmann et al. (2007)</td>
</tr>
</tbody>
</table>
Chapter 4

Risk-adjusted Return on Capital

4.1 Introduction

Markowitz (1959) introduced with the portfolio theory a simple framework to model the trade-off between risk and return and exhibited how an investor would be able to find the best mix of risky investments in his portfolio according to his risk preferences. This affected the optimal allocation of capital to risky investments, as an investor can form optimal portfolios by maximizing a RAPM that is consistent with the CAPM.

Financial institutions apply a similar approach for allocating risk capital to risky business activities, as part of a risk-adjusted performance measurement system. Such a system is necessary to ensure that the decisions of a financial institution reflect the interests of its stakeholders, i.e. shareholders and bondholders. Consequently, financial institutions should base their capital allocation process on some RAPMs that determine the ideal mix of risky business activities, keeping in mind the objective to maximize the value for its shareholders. This is in fact equivalent to the allocation problem of an optimal investment portfolio, but with additional constraints. The optimal allocation of risk capital to different business lines depends, for example, on factors such as the correlation of different business lines and regulatory restrictions.

Alexander (2008b) stated that an optimal capital allocation strategy for a financial institution can be achieved by the following three steps:

1. Assess the risk capital for all risky business activities.
2. Aggregate the risk capital across the risky business activities, taking their dependencies into account.
3. Adjust the capital allocation to the different business activities in a way that maximizes the applied RAPM, given certain risk capital constraints.

The first step refers to the quantification of risk exposure within all risky business activities in financial institutions. The second step relates to the diversification effect when aggregating risk, as not all risks are identically correlated. The final step refers to the fact that once a financial
institution has assessed its total risk exposure, the institution can determine the amount of Economic Capital (EC) needed to justify its target financial strength. Subsequently, the capital should be allocated such that it reflects the individual business activities contribution to the overall risk exposure of a bank. This can be achieved by maximizing an applied RAPM.

Financial institutions and in particular banks usually apply the risk-adjusted return on capital (RAROC) as their preferred RAPM for such a capital allocation. As a result the capital budgeting process and therefore the decision of whether to invest in a business line or a project is usually based on the RAROC. If, for instance, a business line is expected to earn a RAROC in excess of an appropriate hurdle rate that reflects the costs of the institution’s equity capital (i.e. risk capital), this business line is deemed to create shareholder value. Consequently, financial institutions should allocate the EC to the business activities that offer the highest RAROC, as this maximizes the value of the shareholders.

The sections of this chapter are structured along the three steps (1), (2) and (3) mentioned above. The first and the second sections introduce the EC and regulatory capital, the third section provides an overview of the common methods of measuring risk. The fourth section introduces methods for aggregating risk and the final section gives an introduction to the RAROC as a tool for optimal economic capital allocation and performance evaluation.

4.2 Economic Capital

Economic Capital is defined as the amount of capital that banks or other financial institutions must hold to limit the probability of default within a given confidence level over a certain time horizon. It is therefore the amount of financial cushion required to absorb severe unexpected losses, whereas unexpected losses are defined as the difference between the actual losses and the expected losses (see figure 4.1). The time horizon is usually one year, while the confidence level depends on the bank’s target financial strength (i.e. credit rating). Ordinarily, EC is set at a confidence level slightly below 100%, as it would be too costly or even impossible for a bank to guarantee to never default - independent of any future loss experience. This implies that there still remains a small probability that losses will exceed the amount of economic capital (Crouhy, Galai & Mark, 2001). As an illustration, a common objective for large international banks is to maintain an Aa credit rating, which implies a one year probability of default of about 0.03%16. According to Hull (2007) EC can also be regarded as a “currency” for the risk-taking within a bank, considering a business unit may only take a certain risk when the appropriate amount of EC for that specific risk is allocated. EC therefore plays a crucial role by helping a bank to price its risks and as well as to set sophisticated risk limits for its individual business units. Moreover, EC is a consistent measure of risk, which converts any quantified risk a bank is exposed to into an amount of capital.

---

16 The rating is according to the standard of Moody’s. See, for instance, Duffie and Singleton, 2003, table 4.2 and Hull (2007) table 11.1.
needed for the risks taken. These risks involve the three, by the regulator well defined, risk categories; credit risk, market risk and operational risk, as well as additional risk categories that may have an influence on a financial institution such as liquidity, reputation and business risks. Thus, the determination of EC and its allocation on the various business units is primarily a tool for risk budgeting, where limits are set by placing upper limits on the EC itself or by increasing the cost of risk capital.

![Economic Capital Diagram](image)

**Figure 4.1:** This figure illustrates the idea behind the economic capital, with a loss density distribution for credit risk (clearly recognizable due to the typical positively skewed distribution). The actual loss can be either smaller or bigger than the expected loss. Assuming that the actual loss is bigger than the expected loss, the difference is called unexpected loss. Economic capital is defined as the difference between the expected loss and some low-quantile of the loss distribution. Hence, EC is the cushion that banks should hold against unexpected losses in order to guarantee solvency. The residual loss potential is the probability that the unexpected losses exceed the EC, which would cause a bank to default. For more detailed information see, for instance, Blum, Overbeck and Wagner (2003).

### 4.3 Regulatory Capital

As discussed in chapter 2, the regulator is deeply interested in a smooth functioning of the financial system and is therefore concerned about an adequate minimum capitalization of banks. This lead to the nomination of the Basel Committee on Banking Supervision (Basel Committee), which formulates broad supervisory standards and guidelines of best practice with the objective that the regulators will implement them in their own national system. Hence, it can be seen as the closest entity to an international banking regulation that has emerged.

In 1999 the Basel Committee declared to introduce a new capital adequacy framework, known as Basel II regulatory framework, to replace the previous regulation framework. The main target of
the Basel II framework was to ensure adequacy of the banks capitalization, as well as to encourage best-practice risk management which support a smooth functioning of the financial system.

Basel II is a comprehensive framework based on a three-pillar concept, where the first pillar contains minimum capital requirements, the second pillar is concerned with the supervisory review process and third pillar deals with the market discipline. These three pillars are considered as equally important. Consequently, if a bank chooses to calculate risk capital with internal models, the supervisors receive more responsibility to validate and approve these models.

The first Pillar is concerned with the maintenance of regulatory capital and encompasses the three major risks that a bank faces, i.e. credit risk, operational risk and market risk. For the assessment of credit risk and operational risk Banks may choose three different approaches of increasing sophistication. For credit risk there are namely, the standardized approach, the foundation internal rating based (IRB) approach and the Advanced IRB approach. For operational risk, the three approaches are the basic indicator approach, the standardized approach and the advanced measurement approach (AMA). In contrast, the market risk remains unchanged to the previous framework that suggests using the VaR measure in calculating\(^{17}\).

As response to the financial crisis, which begun in mid 2007, the Basel Committee fundamentally supplemented the Basel II framework. In particular, the most recent revisions encompass the Guidelines for computing capital for incremental risk in the trading book (2009), the Revisions to the Basel II market risk framework (2009), the Enhancements to the Basel II framework (2009), the International framework for liquidity risk measurement, standards and monitoring (2009) and the proposal to Strengthening the resilience of the banking sector (2009).

They, in general, cover the following five aspects: (1) They raise the quality, consistency and transparency of the capital base in order to ensure that the banking system can better absorb losses. (2) They introduce a global minimum liquidity standard for international active banks, so that banks maintain an adequate level of high quality assets that can be converted into cash to meet its liquidity needs for a 30 day horizon. (3) They introduce a leverage ratio as a supplementary measure to the existing framework of the first pillar, which should prevent the build-up of excessive leverage in the banking system. (4) They strengthen the capital requirements for counterparty credit risk exposures from derivatives, repos and securities. (5) Finally they introduce a countercyclical capital buffer that builds up capital buffer in good times for periods of stress.

All these changes were proposed with the same objective, to increase the capital cushion allowing banks to withstand higher losses over longer periods of extreme volatility in financial markets with the hope there will not be another necessity for large scale government interventions.

\(^{17}\) The readers interested in a detailed overview of the evolution of the Basel framework are referred to Crouhy et al. (2005) or Alexander (2008b).
4.4 Methods of Measuring Risk Capital

In contrast to the calculation of the regulatory capital, where risk managers are not free to estimate risk with an approach they perceive to be most appropriate, the EC estimation can be based on any internal risk model or risk measure they prefer, as long as the methodology is accepted by the board of directors. In practice the most often used approach to estimate EC is the so-called bottom-up approach. This approach estimates the risk exposure for the different risk types in the business units and aggregates it to the risk of business units and further to the total risk exposure for the entire bank. Once a financial institution has assessed its total risk exposure, the institution can determine the amount of EC needed to justify its target credit rating (see Alexander, 2008b).

In general, banks estimate EC based on the quantification of extreme losses, as they tend to mimic the Basel II metrology. As a result EC is usually calculated based on either a scenario stress tests alone, down side risk measures such as Value-at-Risk (VaR) and Expected Shortfall (ES) or - as suggested by the Basel Committee (2009) for market risk - a mixture of VaR or ES and scenario stress tests (i.e. in the case of market risk, the stressed Value-at-Risk). The subsequent sections will introduce these three, most common in practice, risk measures, namely the Value-at-Risk, the Expected Shortfall and the scenario stress tests, in more detail.

4.4.1 Value-at-Risk

The Value-at-Risk (VaR) is the downside risk measure that has probably received most attention in the banking industry. It was introduced by JP Morgan in response to the financial disasters of the early 1990s, as a method that quantifies possible expected losses due to extreme movements of the underlying risk factors.

For a given time horizon $\Delta$ and confidence level $\alpha \in (0,1)$, the VaR is defined as the maximum loss of a security or portfolio that is not exceeded with probability of $\alpha$, i.e.

$$VaR_\alpha = q_\alpha,$$

where $q_\alpha$ is the relevant quantile of the loss distribution. Typical values for $\alpha$ are $\alpha = 0.95$ and $\alpha = 0.99$. The a time horizon $\Delta$ for market risk is usually 1 or 10 days and for credit and operational risk usually one year (see McNeil et al., 2005).

When calculating the VaR, it is necessary to assume that the risk exposure remains constant over the time horizon in which the VaR is computed. As a result, as the time horizon extends, the VaR loses partly loses its significance. The time horizon is usually chosen in such a way that it corresponds to the liquidity of an asset, therefore, the more illiquid an asset the longer the time horizon. Moreover, the underlying model is assumed to reflect exactly the future distribution of the

---

18 Although the application is in reference banks, it could also be extended to an insurer or other financial institutions that face the similar risks (Rosenberg and Schuermann, 2006).
uncertain risk factors. Standard methods used in the financial industry for calculating the VaR are, the variance-covariance method, the historical simulation and the Monte Carlo simulation. While both the variance-covariance and the Monte Carlo simulation deduce the interaction of the risk factors from the past into the future, the historical simulation directly applies past distribution to the forward looking risk horizon. Thus, all three methods are based on the assumption that historical data can be used for predicting the future, which many academics consider as controversial.

![Figure 4.2](image)

**Figure 4.2:** This is a representation of the market risk of a portfolio that invested $100,000 in the S&P 500 using the one-day VaR at a confidence level of $\alpha = 0.99$. While the VaR expects – over the whole sample of 2516 data points (11 years) – the loss to exceed the VaR ~ 25 times, the VaR was in reality 50 times exceeded. Accordingly, it illustrates that the normality assumption usually is indeed violated in practice.

However, Alexander (2008b) states that even if historical data seems fairly poor in predicting the future over a risk horizon longer than a few months, experiences have proven that in the absence of a shock, market behaviors are unlikely to alter completely over shorter risk horizons. It is therefore reasonable to base short-term VaR on estimations of historical data, but as the risk horizon increases other approaches might be more accurate, such as stress scenarios.

The application of the VaR as measure for EC has also been fundamentally criticized due to the fact that it has poor aggregation properties (i.e. weaknesses mapping diversification effects). This stems from the fact, that it does not satisfy the properties of coherence, and in particular that the VaR is not subadditive (Artzner et al., 1999; Acerbi, 2003). The lack of subadditivity can lead to unexpected and undesirable outcomes, as it can happen that a portfolio has a higher VaR, even though it is more diversified than another portfolio with a lower VaR.

---

19 A risk measure $p$ is subadditive if it satisfies $p(X + Y) \leq p(X) + p(Y)$, for any random variables $X$ and $Y$ in its range of definition.
Another drawback relates to the fact that the VaR does not provide any information about the amount of possible losses that occur with a probability lower than $1 - \alpha$. This is, in particular, relevant for business lines that have a low VaR, but face significant losses in case of operational failures.

### 4.4.2 Expected Shortfall

Expected Shortfall (ES), sometimes also referred to as Conditional Value-at-Risk (CVaR), is another approach to estimate risk (Acerbi & Tasche, 2002). In fact, the ES is today even preferred over VaR by most risk managers, as it fulfills the property of subadditivity. Hence, unlike the VaR, the ES always takes the diversification effect into account (see Artzner et al., 1999). Assuming that the loss distribution is continuous, the ES is given as

$$ES_\alpha = \frac{1}{1 - \alpha} \int_\alpha^1 q_d dp.$$  

(4.2)

The ES only depends on loss distribution and is always $ES_\alpha \geq VaR_\alpha$. The ES measure can be interpreted as the average of quantiles, where tail quantiles $p \geq \alpha$ have an equal weight and the non-tail quantiles $p < \alpha$ have a weight of zero.20 When the ES is applied as a measure for quantifying the Economic Capital, a financial institution anticipates to remain solvent up to the expected loss of a worst case, i.e. the average loss of possible losses that occur with a probability lower than $1 - \alpha$. Therefore the link between the confidence level $\alpha$ and the insolvency is omitted. In fact, it is more intuitive to associate negative outcomes of risk-taking with the insolvency of an institution, rather than with the expected loss in the event of a worst case.

### 4.4.3 Scenario Stress Tests

Scenario stress tests unlike the VaR and ES are not based on normal market conditions, as they are a tool for quantifying the size of potential losses under stress events. Their ultimate purpose is to simulate stress events similar to the ones that have occurred during the financial crisis where multiple things went wrong at the same time. A stress event can therefore be defined as an exceptional but a still possible event in the market to which a portfolio is exposed.

As a response to the weaknesses of the stress testing practice within banks, revealed by the financial crisis, the Basel Committee tried to strengthen practice of stress testing with the publication of the *Principles for sound stress testing practices and supervision* in May 2009. They claimed that stress tests are not only a remedy for all weaknesses of risk management, but also

---

20 This implies that if a user applies this risk measure, he is assumed to be risk-neutral between better and worse tail outcomes, as the measure weighs tail losses equally.
highlighted the crucial role of stress testing as a supplement for other risk management approaches. The Basel Committee has, moreover, emphasized the importance of stress testing in determining capital adequacy of banks by requiring banks to calculate a stressed VaR for minimum regulatory capital for market risk (see the Revision to the Basel II market risk framework, 2009).

A possible definition of scenario stress tests is given by McNeil et al. (2005), where the portfolio of risky securities is given as $l_{[t]}$, which is the loss operator of the portfolio. A set of risk factors changes (i.e. scenarios) is given by $X = \{x_1, \ldots, x_n \}$ and a vector of weights is given by $w = (w_1, \ldots, w_n)' \in [0,1]^n$. Both the risk factors and the weights need to be determined by the user. The risk exposure of the portfolio is then given by

$$\psi[X, w] = \max \{w_1 l_{[t]}(x_1), \ldots, w_n l_{[t]}(x_n)\}. \quad (4.3)$$

Thus, the exposure of a portfolio is measured as the maximum loss of the portfolio taking all scenarios into consideration. Alexander (2008b) emphasizes that despite the fact that such an approach is commonly used in practice, there is no standard approach for stress testing. As a result various approaches of stress testing have been developed, which are usually categorized by the type of change they consider and also by the data used to derive them. Common stress testing approaches are the VaR based on stressed covariance matrices and the stress tests based on principal component analysis.

### 4.5 Aggregation of Risk

When a financial institution applies the bottom-up approach for assessing its total risk exposure, it is required to assess all relevant business risks and quantify those risks with risk measures such as the VaR or ES. The relevant risks encompass the three risk categories well defined by the regulator: market risk, credit risk and operational risk, as well as any other risk that might have a significant impact on an institution. In practice financial institutions usually calculate market, credit and operational risk loss distributions for their different business units. To arrive at the total risk exposure of a financial institution, they are required to aggregate these loss distributions.

Rosenberg and Schuermann (2006) name three simple methods for aggregating these loss distributions, which are namely the additive, the normal, and hybrid approach. These approaches are, however, only valid if some critical assumptions concerning the relationship between individual distribution quantiles and portfolio quantiles are satisfied. Thus, the quantiles of the individual standardized returns are assumed to be the same for the portfolio returns, which, for example, is satisfied by the family of the elliptical distributions.

---

21 Before aggregating these loss distributions it is crucial to equalize them by the time horizon and the confidence level (see The Joint Forum, 2003)
(1) The additive approach, moreover, restricts the correlations to be perfect, but does not restrict the quantiles. This is a very conservative approach, as it assumes a perfect correlation of the individual risks (i.e. a correlation of one) and therefore ignores a possible diversification across business activities. In reality, however, risks do usually not add up arithmetically, so that the total risk is less than the sum of individual risks. This was proven by Rosenberg and Schuermann (2006), who showed that this approach in general overestimates risk by 20 to 35 percent.

(2) The normal approach, in contrast, assumes that the quantiles come from a normal density, but allows for estimated correlations. This usually underestimates the capital requirements, as the normal assumption does not take into account higher moments of a distribution (i.e. skewness and kurtosis). In fact, Rosenberg and Schuermann (2006) showed that the shapes of the loss distributions of risk vary considerably. While, for instance, market risk typically has a loss distribution that is nearly symmetrical and approximated normal, credit and in particular operational risk generate more skewed and heavy tailed distributions due to the occasional occurrence of extreme losses. As a result, the normal assumption has shown to be indeed inappropriate in aggregating such distributions.

(3) The hybrid approach is the least restrictive of these approaches, because it allows the correlations and the marginal quantiles to depend on estimations. Rosenberg and Schuermann (2006) show that the hybrid approach provides good approximations even when the loss distributions are non-normally distributed.

Assuming that the relation between the individual distribution quantiles and the portfolio quantiles is not satisfied, more sophisticated techniques might be required, such as copula approach. The idea behind the copula approach is that a joint distribution can be factored into the marginal distributions and a dependence function called copula. As the term copula indicates the copula, couples the marginal distributions together creating a joint distribution, where the dependence relationship is entirely determined by the copula, and scaling and shape are fully determined by the marginal distributions. As emphasized by Rosenberg and Schuermann (2006), copulas are especially valuable for combining risks when the marginal distributions are estimated individually. As a result, they are perfectly suited for the aggregation of loss distributions (see McNeil, 2005; Hull, 2007)\(^{22}\).

\(^{22}\) It should be noted, that these aggregations usually cause high model risk, as an inappropriate assumption about the dependencies is likely to be the most important source of model risk in financial institutions that apply such approaches for risk assessment.
4.6 RAROC

Once a financial institution has assessed its total risk exposure, the institution can determine the amount of EC needed to justify its target credit rating. The EC should subsequently be allocated to the business activities, based on their risk contribution to the institution’s overall risk exposure. In order to ensure that the EC is allocated in a way that maximizes the shareholder value, banks usually apply some RAPMs. The best known RAPM is probably the risk-adjusted return on capital (RAROC), which was developed in the late 1970s by a group at Bankers Trust. Its original intent was to determine a bank’s required equity capital in order to limit its credit risk exposure. However, today the RAROC system is rather a tool for steering an entire financial institution, through efficient capital allocation.

Zaik et al. (1996) mention two basic reasons as to why such an allocation is crucial for a bank: First, the overriding objective of an efficient capital allocation is to determine a bank’s optimal capital structure, i.e. the optimal debt-equity ratio in order to minimize the overall funding costs. Second, the RAROC framework evaluates business unit’s contribution to shareholder value and reveals the business units that destroy shareholder value. It therefore provides a basis for effective capital budgeting and incentive compensation within an institution.

The RAROC is a performance measure similar to any other RAPM, while it also compares the trade-off between risk and reward. Crouhy et al. (2005) define RAROC as:

$$RAROC = \frac{\text{After Tax Expected Risk Adjusted Net Income}}{\text{Economic Capital}}.$$  

(4.4)

The RAROC makes also an adjustment in the numerator with the subtraction of a risk factor from the expected net income (e.g. the expected loss). In the case of the expected loss, the subtraction is necessary as these costs are – like any other business costs – already priced into the business transactions, therefore no EC must be set aside for this risk.

If a bank wants to maximize its shareholder value, it is required to realize all new activities that exceed the cost of capital. This, however, might be inefficient. If executive managers are rewarded solely on the basis of RAROC, they are likely to reject value-increasing projects that are supposed to lower their average returns.

To mitigate such a behavior, the so-called economic profit of new activities has to be calculated, which is the RAROC minus a charge for the cost of equity (i.e., shareholder’s minimum required rate

---

23 The interested readers in a unifying framework for allocating the aggregated capital of a financial institution to its business units is referred to Dhaene et al. (2009).

24 The precise definition of RAROC varies across financial institutions, as its exact calculation is usually determined by the executive management in association with the shareholders and the board of directors. However, even though the RAROC is the most common RAPM in the banking industry, there are various other performance measures such as the risk-adjusted return on assets (RAROA) and the return on risk-adjusted assets (RORAA). See Matten (2000) for a comprehensive list of all these RAPMs.
of return), also known as the hurdle rate. The hurdle rate can be determined by applying a pricing model such as the CAPM, which defines the hurdle rate as

$$HR = R_f + \beta_E (E(R_m) - R_f),$$

(4.5)

where $R_m$ is the expected rate of return on the market portfolio, $R_f$ the risk free interest rate and $\beta_E$ the beta of the equity of the firm. As a result, a business activity with a RAROC exceeding the HR is deemed to create shareholder value, whereas a RAROC that falls below the HR is deemed to destroy shareholder value.

In essence, the RAROC-based capital allocation can be viewed as an internal capital market, where all business units of a financial institution compete with each other for the scarce resource risk capital (i.e. EC) with the main objective to maximize shareholder value. Thus, the RAROC framework ultimately provides information about where shareholder value is created and where it is destroyed to the executive management of a financial institution\textsuperscript{25}.

\textsuperscript{25} The reader interested in a mathematical framework for optimal capital allocation is referred to the Euler capital allocation principle, which requires a risk measures to be compatible with the RAROC and adds up to portfolio-wide risk. The Euler allocation is the process of allocating capital to assets and portfolios with the calculation of the Euler contribution (see Tasche, 2008).
Chapter 5

Empirical Example

5.1 Data

The examined data involves 6 return data series with 210 data points\(^{26}\). Table 5.1 shows the main characteristics of the sample. It provides the mean return, the maximum return, the minimum return, the standard deviation, the skewness, the kurtosis and the Jarque-Bera value for each data series. The Jarque-Bera test reveals that at a confidence level of 95%, the normal assumption cannot be rejected for data series E0, E1 and E2. In contrast, data series F0, F1 and F2 reject the normal assumption even at a 99% confidence level.

Figure 5.1 provides a graphical representation of the data series, where the histograms appear on the y-axis. It is noticeable that most of the density of the series F0, F1 and F2 is centered around a value of \(~0.07\), while no value falls below zero.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>E0</th>
<th>E1</th>
<th>E2</th>
<th>F0</th>
<th>F1</th>
<th>F2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0082</td>
<td>0.0095</td>
<td>0.0104</td>
<td>0.0081</td>
<td>0.0099</td>
<td>0.0105</td>
</tr>
<tr>
<td>Median</td>
<td>0.0090</td>
<td>0.0093</td>
<td>0.0109</td>
<td>0.0068</td>
<td>0.0079</td>
<td>0.0087</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.0194</td>
<td>0.0238</td>
<td>0.0242</td>
<td>0.0194</td>
<td>0.0238</td>
<td>0.0242</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.0032</td>
<td>-0.0072</td>
<td>-0.0070</td>
<td>0.0018</td>
<td>0.0018</td>
<td>0.0018</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0052</td>
<td>0.0077</td>
<td>0.0078</td>
<td>0.0039</td>
<td>0.0062</td>
<td>0.0063</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.3777</td>
<td>0.0474</td>
<td>-0.1499</td>
<td>1.0424</td>
<td>0.9176</td>
<td>0.8517</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.3134</td>
<td>-0.7238</td>
<td>-0.6767</td>
<td>0.1738</td>
<td>-0.2812</td>
<td>-0.4733</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>5.9021</td>
<td>4.8062</td>
<td>4.9303</td>
<td>37.6623</td>
<td>29.8531</td>
<td>27.1316</td>
</tr>
<tr>
<td>Probability</td>
<td>0.0523</td>
<td>0.0904</td>
<td>0.0850</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

\(^{26}\) Data received March 24, 2010, from Prof. Dr. Karl Frauendorfer.
5.2 Rankings and Comparison

This section compares the six data series presented in section 5.1, by applying the most popular RAPMs that have been introduced in chapter 3. The calculations are based on the assumption that the data points correspond to monthly returns. The risk-free-rate is calculated according to the 1 month Libor of 0.40% p.a. (i.e.~0.03% p.m.). The threshold is chosen at a value \( \tau = 0.003 \) in general and \( \tau = 0.007 \) for the RAPMs market with a star*.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>SR</th>
<th>GSR</th>
<th>MPPM (2)</th>
<th>MPPM(10)</th>
<th>MPPM(100)</th>
<th>Kappa(4)</th>
<th>Omega</th>
</tr>
</thead>
<tbody>
<tr>
<td>E0</td>
<td>1.5096</td>
<td>4</td>
<td>0.0493</td>
<td>0.0480</td>
<td>0.0329</td>
<td>1.8124</td>
<td>9.3649</td>
</tr>
<tr>
<td>E1</td>
<td>1.1901</td>
<td>6</td>
<td>0.0644</td>
<td>0.0616</td>
<td>0.0320</td>
<td>1.6495</td>
<td>8.4580</td>
</tr>
<tr>
<td>E2</td>
<td>1.3011</td>
<td>5</td>
<td>0.0759</td>
<td>0.0731</td>
<td>0.0412</td>
<td>1.9397</td>
<td>9.7527</td>
</tr>
<tr>
<td>F0</td>
<td>1.9646</td>
<td>1</td>
<td>0.0484</td>
<td>0.0477</td>
<td>0.0406</td>
<td>15.6508</td>
<td>864.3776</td>
</tr>
<tr>
<td>F1</td>
<td>1.5593</td>
<td>3</td>
<td>0.0704</td>
<td>0.0687</td>
<td>0.0526</td>
<td>16.7686</td>
<td>253.4702</td>
</tr>
<tr>
<td>F2</td>
<td>1.6110</td>
<td>2</td>
<td>0.0769</td>
<td>0.0751</td>
<td>0.0582</td>
<td>18.4001</td>
<td>277.7286</td>
</tr>
</tbody>
</table>

Table 5.2 compares the rankings of the Sharpe ratio (SR), the Generalized Sharpe ratio (GSR), the Manipulation Proof Performance Measure (MPPM(2)) with a CRRA of 2, the MPPM(10) with a CRRA of 10, the MPPM(100) with a CRRA of 100, the Kappa(4) index and the Omega performance measure. The SR and the GSR provide identical rankings that also corresponds more or less to the rankings of the Kappa(4) index and the Omega. In contrast, the MPPM with CRRA of 2 and 10 rank series F0 as 6th and therefore least favorable return distribution series, whereas series F0 is at first position according to SR, the GSR and Omega.
The observed inconsistency of the rankings mainly stems from the low standard deviation of the data series F0. The MPPM and Kappa(4) are not as sensitive to the standard deviation as the rest of the RAPMs. All RAPMs rank data series F2 either on first or on second place, however.

The accuracy of the MPPM is illustrated by the fact that the ranking of the MPPM changes, for an extremely risk-averse individual. The ranking of data series E2 moves down from the second place to the third, while data series F1 moves up from the third place to the second place. This relative change is caused by the higher probability of lower returns of E2 compared to F1. As long as an individual is not highly risk-averse, the individual would weigh the potential higher returns higher.

Table 5.3: Ranking of Kappa and Omega

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Omega</th>
<th>Kappa(1)</th>
<th>Kappa(2)</th>
<th>Kappa(3)</th>
<th>Kappa(4)</th>
<th>Kappa(1)*</th>
<th>Kappa(2)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>E0</td>
<td>9.3649</td>
<td>8.3649</td>
<td>3.1445</td>
<td>2.1989</td>
<td>1.8124</td>
<td>72.43%</td>
<td>0.3483</td>
</tr>
<tr>
<td>E1</td>
<td>8.4580</td>
<td>7.4580</td>
<td>2.8973</td>
<td>2.0227</td>
<td>1.6495</td>
<td>120.31%</td>
<td>0.6125</td>
</tr>
<tr>
<td>E2</td>
<td>9.7527</td>
<td>8.7527</td>
<td>3.3664</td>
<td>2.3657</td>
<td>1.9397</td>
<td>196.16%</td>
<td>0.8875</td>
</tr>
<tr>
<td>F0</td>
<td>864.3776</td>
<td>863.3776</td>
<td>59.5787</td>
<td>24.4374</td>
<td>15.6508</td>
<td>113.93%</td>
<td>0.7078</td>
</tr>
<tr>
<td>F1</td>
<td>253.4702</td>
<td>253.4702</td>
<td>42.7387</td>
<td>23.0417</td>
<td>16.7686</td>
<td>325.28%</td>
<td>1.7233</td>
</tr>
<tr>
<td>F2</td>
<td>277.7286</td>
<td>277.7286</td>
<td>46.9419</td>
<td>25.2827</td>
<td>18.4001</td>
<td>457.23%</td>
<td>2.2511</td>
</tr>
</tbody>
</table>

Table 5.3 compares all Kappa Indices (1 to 4) and Omega (which is similar to Kappa(1) – 1) to each other. Similar to the comparison in table 5.2, the biggest difference is related to data series F0. Higher order Kappa indices show to be less sensitive to the standard deviation, compared to lower order ones. However, the rankings are very similar; in particular all Kappa indices with the same threshold rank data series E0, E1 and E2 identically. Table 5.3, moreover, shows that the Kappa indices are – as stressed in chapter 3 – indeed very sensitive to the chosen threshold. For instance, the rankings from Kappa(1) to Kappa(1)* substantially vary. The sensitivities of the Kappa indices to different thresholds are also shown in figure A.1 in the appendix.
Chapter 6

Outlook and Conclusion

6.1 Outlook: Spectral Risk Measures

So far, spectral risk measures have only been used to measure risk, independently of returns. However, these measures have the potential to be used as denominator in RAPMs in the future. One of the newest innovations in the risk manager’s toolbox are the so-called spectral risk measures proposed by Acerbi (2002, 2003).

Even though spectral risk measures are quantile-based risk measures, they are able to directly refer to a user’s risk spectrum and risk aversion function. More formally, spectral risk measures attach different weights to the \(i\)-th quantile of a distribution. Therefore, they are coherent as long as they attach at least can equal, i.e. non-decreasing, weight to all quantiles above the Value-at-Risk (VaR).

Spectral risk measures are defined as risk measures \(M_\phi\) that weigh the averages of the quantiles \(q_p\) as

\[
M_\phi = \int_0^1 \phi(p)q_p dp ,
\]

where the weighting function, \(\phi(p)\), is known as the risk spectrum or risk-aversion function and still needs to be determined. Spectral risk measures include both, the Value-at-Risk (VaR) and the Expected Shortfall (ES) as special cases. The VaR implies a \(\phi(p)\) function that attaches \(p = \alpha\) an infinite weight, and every other outcome, a zero weight. The ES implies a discontinuous \(\phi(p)\) that takes the value of 0 for profits or small losses and takes a constant value for high losses. These are, however, not “well behaved” spectral risk measures, as they are inconsistent with risk-aversion. Risk-aversion implies that the weights rise smoothly: the more risk-averse an individual, the faster the weight will rise.

The conditions that \(\phi(p)\) must satisfy for \(M_\phi\) to become coherent are:

- **Non-negativity**: \(\phi(p) \geq 0\) for all \(p\) in the range \([0,1]\).
• **Normalization:** \( \int_{0}^{1} \phi(p) \, dp = 1. \)

• **Weakly increasing:** \( \phi(p_1) \leq \phi(p_2) \) for all \( 0 \leq p_1 \leq p_2 \leq 1. \)

Condition 1 requires the weights to be non-negative, and the second requires the sum of the probability weight to be 1. The third condition directly reflects an individual's risk-aversion, since it requires that bigger weights need to be attached to higher losses, or at least not lower weights than that have been attached to the lower loss.

In other words, spectral risk measures can be used to attach different weights to the \( i \)-th quantile of a loss distribution. They are coherent risk measures as long as the weightings that each quantile receives are a non-decreasing function of the quantile itself. Thus, spectral risk measures attach higher weights to higher losses and therefore directly reflect an investor’s risk aversion.

**Figure 6.1:** This is a comparison of a coherent and a non–coherent risk aversion function. The function on the right is not subadditive and therefore not coherent, as it does not always attach non-decreasing weight to higher losses. The function on the left is subadditive and therefore coherent, as it attaches higher weights to higher losses (Acerbi, 2003, p. 36).

In combination with returns, spectral risk measures could be used as a RAPM to assess allocation problems. On the one hand, the spectral risk measures could be used in the RAROC as a risk measure for economic capital (EC), given that it is a quantile-based risk measure and therefore adequate for setting risk limits. Thereby, the risk that a corporation will face insolvency can be reduced, as the VaR usually underestimates risk (see, also figure 4.2).

On the other hand, spectral risk measures could be applied as the risk measure in the RAPM ratio used for optimal investment selection decisions, as it is able to reflect exactly the risk-aversion of
individual investors and therefore is also consistent with the utility theory, unlike most other downside risk measures.

### 6.2 Conclusion

This thesis has analyzed the broad existing literature on risk-adjusted performance measures (RAPMs) used for asset and capital allocation in financial institutions. In contrast to the existing financial literature, which usually treats these fields separately, this thesis considers both the application areas of RAPMs.

In particular, for the area of asset allocation, a wide range of new RAPMs have recently been developed. These different RAPMs can be clustered into four categories, grouped based on their similar mathematical characteristics: Mean-variance RAPMs, CAPM RAPMs, downside risk RAPMs and preference-based performance measures. The mean-variance performance measures still enjoy great popularity in practice even though it is the oldest of these categories. In particular the Sharpe ratio (SR), which is defined as the expected excess return divided by standard deviation of those returns, is widely used in practice. The CAPM performance measures were introduced and became popular along with the development of the CAPM. These performance measure, however, are based on very restrictive assumptions, which when violated usually lead to fallacious results. Hence, these performance measures are accurate as long as investors are only concerned about systematic risk.

The downside risk performance measures have been engineered to resolve the issue that mean-variance and CAPM performance measures do not take into account higher moments of a return distribution, which is typically the case for returns of hedge funds. A drawback, however, is that these performance measures often have no plausible relation to risk-averse utility functions and therefore do not provide information about preferences from their ranking. As an answer to this drawback, academics and practitioners developed the preference-based performance measures, the newest RAPMs in the toolbox of asset managers. They are derived from utility functions characterizing investor’s preferences to higher moments. They have proven to be particularly useful when return distributions are odd-shaped, as they are not limited to a specific distribution shape.

The analysis of RAPMs sensitivity to manipulation has shown that most RAPMs can be subject to manipulation activities that are used to alter the performance of an applied RAPM. Solely the preference-based measures have proven to be resistant to most of these activities. However, not only manipulations may influence the rankings, but also the choice of the RAPM itself. This fact has been demonstrated by the application of empirical data for the most popular RAPMs.

Overall, the thesis concludes that the Sharpe ratio (SR) provides on average fairly good rankings, even if return distributions are skewed and heavy tailed. The performance of traditional RAPMs – and in particular the SR – is, however, prone to high positive skewness and excess kurtosis. This stems from the fact that the variance is a symmetric risk measure, punishing positive and negative
deviations from the mean. As a result, the performance is usually undervalued. Nevertheless, investors are still well-advised to apply the SR as a RAPM for selecting alternative investment opportunities, as it is easy to calculate and provides surprisingly adequate rankings even in the presence of non-normally distributed returns.

In contrast, if senior managers or investors assess the performance of asset managers, a more sophisticated RAPM might be more appropriate given that the SR has shown to be prone to manipulations. The best suited RAPM for such an evaluation is the MPPM, which has proven to be resistant to any manipulation attempts by asset managers.

In the area of capital allocation, the risk-adjusted return on capital (RAROC) is the predominant RAPM. This RAPM was developed to determine a bank’s appropriate level of equity capital in order to limit its credit risk exposure. Today, RAROC is used as a guide for optimal capital allocation within financial institutions. Its ultimate purpose is to evaluate the contribution of each business unit to shareholder value. Therefore, RAROC provides information to the executive management of an institution over where value is created and where it is destroyed. It is an adequate measure for efficient capital budgeting and incentive compensation in financial institutions.

Advancing forward, the development of new measures for risk-adjusted performance evaluation is likely to continue. One potential enhancement might be the application of spectral risk measures for existing RAPMs. With their ideal mathematical characteristics of relating directly to an individual’s risk-aversion function and being a quantile-based risk measure, spectral risk measures also have the potential to further improve the accuracy of risk adjusted performance measurement in the area of both asset and capital allocation.
References


The MPPM is applying the following power utility function

\[ U(W) = \frac{W^{1-p}}{(1-p)}, \quad p > 1. \]  \hspace{1cm} (A.1)

The relative risk aversion (RRA) is given by

\[ \text{RRA} = -W \frac{U''(W)}{U'(W)}. \]  \hspace{1cm} (A.2)

Substituting equation (A.1) into equation (A.2) and rearranging results in

\[ \text{RRA} = -W \frac{-p \cdot W^{-p-1}}{W^{-p}} = p. \]  \hspace{1cm} (A.3)

Consequently, the RRA is constant, which corresponds to a constant relative risk aversion (CRRA).
8.2 Appendix to Chapter 5

Figure A.1: Rankings based on Kappa Indices

27 The thresholds appear on the x-axis and the kappa values on the y-axis.
8.3 Declaration of Authorship

I hereby declare

- that I have written this thesis without any help from others and without the use of documents and aids other than those stated above,

- that I have mentioned all used sources and that I have cited them correctly according to established academic citation rules,

- that I shall not pass on any copies of this thesis to any third parties without the president’s consent, with the exception of fellow students or persons who have provided me with essential information for this thesis, to whom I may pass on copies of this thesis after the procedure has been concluded.

Brunegg, May 25, 2010

Philipp Schmid