D I P L O M A R B E I T

Entropy-Based Portfolio Selection

ausgeführt am

Institut für
Stochastik und Wirtschaftsmathematik
TU Wien

unter der Anleitung von

Ao.Univ.Prof. Dipl.-Ing. Dr.techn.
Klaus Felsenstein

durch

Lukas Meisinger
[redacted]

Wien, am 22. Oktober 2024

Kurzfassung

Diese Diplomarbeit befasst sich mit der Verbesserung des Markowitz-Modells für das Erstellen eines optimalen Portfolios. Das Ziel eines ausgewogenen, diversifizierten und risikoeffizienten Portfolios steht dabei im Vordergrund. Die Schwachstellen ebenjener Methodik sollen durch die Erweiterung um die Entropie ausgemerzt werden. Diese wird zunächst statisch eingeführt. Um die Veränderungen und Schwankungen auf dem Kapitalmarkt jedoch ebenso einfließen zu lassen, wird danach eine Methodik vorgeschlagen, die den Einfluss der Entropie dynamisch über die Zeit gewichtet.

Des Weiteren wird untersucht, ob solche Methodiken zur Optimierung von Portfolios gegenüber vordefinierten Benchmarks führen können. Dazu wird auch die Kullback-Leibler Divergenz verwendet um zu sehen, ob das tracken eines statischen Sets von Assetklassengewichten statt der Gleichgewichtung, die durch die Entropie-Methodik verfolgt wird, zu besseren Ergebnissen führt.

Abschließend wird noch versucht den dynamischen Einfluss der Entropie anzupassen um ein robusteres und stabiles Portfolio zu erlangen. Dazu wird der initiale Versuch, den Parameter empirisch zu steuern, abgelöst von einer Performance-basierten Variante.

Als Resultat hat sich ergeben, dass sich diese Methodiken durchaus eignen, um profitable und zugleich risikoeffiziente Portfolios zu erzeugen. Je nach Art der Benchmark oder des Investmentziels, kann es zu etwas unterschiedlichen Ergebnissen kommen. Während sich die reine Entropie-Methode bei einer konservativeren Benchmark als effizienter erwiesen hat, führt das adaptierte Entropie-Modell bei dynamischeren Zielen zu besseren Ergebnissen.

Wichtig für einen Investor ist, diese Optimierungsmodelle zu verstehen und in Investmententscheidungen mit einfließen zu lassen. Solch quantitave Verfahren können in Hinsicht auf sowohl Risiko als auch Ertrag eines Portfolios einen großen Mehrwert beisteuern.

Abstract

This thesis deals with the improvement of the Markowitz model for the process of creating an optimal portfolio. The goal of a balanced, diversified and risk-efficient one is at the forefront. The weaknesses of Markowitz' methodology are to be eliminated by adding entropy. This is first introduced statically. However, in order to incorporate the changes and fluctuations on the capital market, a methodology is then proposed that dynamically incorporates the influence of entropy over time.

Furthermore, it is examined whether such methodologies can be used for the optimization of portfolios over predefined benchmarks. For this purpose, the Kullback-Leibler divergence is used to see whether tracking a static set of asset class weights instead of an equalibrium, which is pursued by the entropy methodology, leads to better results.

Finally, an attempt is made to adjust the dynamic influence of entropy to achieve a more robust and stable portfolio. For this purpose, the initial attempt to empirically calculate the parameter is replaced by a performance-based variant.

As a result, it has been shown that these methods are certainly suitable for creating profitable and and risk-efficient portfolios at the same time. Depending on the type of benchmark or what investment objective is pursued, the results may vary somewhat. Whereas the pure entropy method has proven to be more efficient with a more conservative benchmark, the adapted entropy model leads to better results with more dynamic targets.

It is important for an investor to understand these optimization models and to incorporate them into investment decisions. Such quantitative procedures can be used to add great value to a portfolio in terms of both risk and return.

Danksagung

Ich möchte mich bei allen Menschen und vor allem Freunden bedanken, die mich sowohl auf meinem Weg durch das Studium sowie beim Verfassen meiner Diplomarbeit begleitet haben. Ein besonderes Dankeschön möchte ich Herrn Professor Klaus Felsenstein aussprechen, der mich fachlich hervorragend unterstützt hat. Desweiteren gebührt mein Dank Herrn DI Patrick Fiedelsberger, der mir bei jedem Schritt meiner Diplomarbeit zur Seite gestanden ist und mich immer motiviert hat. Darüber hinaus möchte ich mich auch bei der Niederösterreichischen Versicherung AG bedanken für die Daten und die Unterstützung, die sie mir zukommen ließen.

Ein besonderer Dank gilt auch meinen Kommilitoninnen Lisa, Steffi und Katrin, ohne deren moralische und organisatorische Hilfe ich das Studium vermutlich nicht geschafft hätte. Des Weiteren möchte ich mich auch bei Alex für das Lesen und Korrigieren meiner Arbeit sowie bei meiner Freundin Nora, dafür dass sie immer an mich geglaubt hat, bedanken.

Abschließend möchte ich noch ein großes Dankeschön an meine Eltern Veronika und Johann sowie an meine Geschwister Lisa und David richten. Die Geduld und Unterstützung, die sie mir während meines gesamten Studiums zukommen ließen, war mehr als Gold wert.

Lukas Meisinger
Ardagger, am 22. Oktober 2024

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am 22. Oktober 2024

Lukas Meisinger

Contents

1 Introduction

In the world of finance, an investor has to navigate through a highly complex universe of different asset classes, securities and financial instruments. Each and every single one of them has their own characteristics. With purchasing an asset comes always the chance for potential gains, but also for possible losses. To get the best out of an investment, it is crucial to understand how one can minimize the risk of losing money and at the same time maximize the return of one's portfolio. However, this is not an easy process since there are numerous ways to measure both of these metrics and beyond that even more options on how to interprete and use them to get the best possible trade-off between risk and return.

The base for the modern portfolio theory as we know it has been laid by Harry M. Markowitz in his infamous "Portfolio Selection" in 1952. According to this article, investors are able to optimize their investment strategies by looking at the trade-off between risk and return. He starts by assuming an investors seeks to get as much expected return as possible while taking as little risk as possible. [1]

However, over the many years a broad range of other optimization models and ideas was created. One of the main problems with the Markowitz model is,that it tends to extreme solutions. That means it overweights certain asset classes while others with not ideal characteristics are excluded or only up to a very small percentage. This contradicts the initial idea of diversification.

To deal with this issue, different authors tried to adjust the model. One of these adjustments is to include the entropy in the objective function. Since the entropy is at its maximum when the assets are equally distributed, this method should avoid such extreme weights in their solution according to Song and Chan ([2]) or Mercurio, Wu and Xie ([3]). As in every other model as well, there are some problems to be addressed. One of them, is that pure entropy would deliver an equally distributed portfolio, ignoring the characteristics of the single assets or asset classes. This might not be very appealing to investors in reality. Therefore, this thesis tries to adress the inclusion of entropy into the Markowitz model while also maintaining a good trade-off to both the risk and the return.

Moreover, the aim of this thesis here is also to try and compare the different models when trying to reach different goals. Insurance companies for instance are financial service providers with a broad palette of various products to sell to their clients with different levels of risk and return. Those products can reach from very conservative investments with a high share in government bonds to more dynamic ones with a bigger equity percentage in the portfolio. Due to this, the optimization objective can vary hugely. Because of this, this thesis tries to compare the models included here with fictive benchmarks to see, which model can produce a portfolio with less risk and more return than these benchmarks without too much deviation from said benchmarks.

2 Distribution of Returns

The assumption made about the distribution of the asset returns is one of the most important ones when it comes to portfolio optimization. Based on the characteristics of the chosen distribution, the inputs, calculation and interpretation can change a lot. The most common ones for asset returns are the normal distribution, the log-normal distribution and the Student's t-distribution, therefore in the following we concentrate on these three.

2.1 Normal distribution

This distribution is a continuous one, which is defined by the two parameters μ, the mean or expectation, and σ^2, the variance. The general form of its density function is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{2.1}$$

The main assumption here is that the returns are symmetrically distributed around the mean, while there exist neither skewness nor fat tails. This means that extreme gains or losses are less probable than in distributions with fat tails, which often leads to underestimating risks since in reality such events are more likely than in this distribution. However, the big advantage of the normal distribution is that the relationship between assets can be fully described by their linear correlation. This leads to a very straightforward modelling of the diversification effect. [4]

2.2 Log-normal Distribution

This distribution is connected to the previously described normal distribution. A random variable is log-normal distributed, when its logarithm is normally distributed, i.e.

$$Y \sim \text{Log-normal}(\mu, \sigma^2) \iff Y = ln(X) \;\; with \; X \sim \mathcal{N}(\mu, \sigma^2) \tag{2.2}$$

Thus, the density has the form

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(ln(x)-\mu)^2}{2\sigma^2}} = \frac{1}{x\sigma}\phi\Big(\frac{ln(x)-\mu}{\sigma}\Big) \tag{2.3}$$

, where $x > 0$, $\sigma > 0$, $\mu \in \mathbb{R}$ and ϕ is the density of standard normal distribution. The resulting distribution is asymmetrical and positively skewed. Consequently, the log-normal distribution has fat tails, which is more adequate for modelling asset returns than using the normal distribution. Additionally, the fact that prices cannot go below zero is also reflected by this distribution. However, using this distribution would lead to increased complexity in the whole portfolio optimization process. The calculation of expected returns, risk and other key figures needs to be adjusted to adress the skewness and kurtosis in this distribution. [5]

2.3 Student's t-distribution

Last but not least is the Student's t-distribution. It is a generalization of the standard normal distribution and like the latter, it is symmetric around zero and has the shape of a bell. The big difference is the former has heavier tails, which are defined by the degrees of freedom $\nu > 0$, which at the same time is the only parameter of the distribution. For large ν the tails begin to get thinner and for $\nu \to \infty$, it becomes the standard normal distribution. Since the Student's t-distribution takes extreme market movements as a more likely event than the normal distribution suggests, it would lead to a more cautious approach in portfolio optimization. The fat tails would result in higher estimated risks and consequently would return in portfolios that are more robust to extreme losses and are more diverse but also more conservative ones.[6]

2.4 Summary

While the log-normal and Student's t-distribution both are closer to the real world than the normal distribution, the latter supplies a mathematically more straightforward way to do a portfolio optimization. The two former would increase the complexity in the whole process exponentially. The calculation of both the portfolio return and the portfolio risk would need to be adapted. The diversification effects might also differ depending on the distribution since the choice of it might also affect the covariance matrix. Therefore, the comparability and interpretation between models with different distributions are more difficult.

Since this thesis aims to examine a portfolio optimization approach including entropy and compare it to the Markowitz model, going forward we will concentrate on the log-normal distribution.

3 Return Measures

The return is a fundamental component of portfolio theory and can be measured in various ways, all of which have their own up and down sides. In most cases, it is not the real returned money one gets from selling an asset, but the growth of the value of an asset or the wealth of an investor while considering different factors. To move forward we look at the most common ones.

3.1 Simple Return

The most straightforward way to measure the growth of the value is by looking at the start value v_0 and end value v_1 after a given period. By dividing the difference between those two by v_0, we reach the percentual growth of value of the asset.

$$R = \frac{v_1 - v_0}{v_0} \tag{3.1}$$

This can be either applied to one asset or to a whole portfolio by taking the sum of all investments. It is the simplest and most basic form to measure the return.

3.2 Total Return

Many assets like bonds or stocks also come with dividends, interests, coupons or other payments one can get by holding onto them. These benefits are taken into account by the total return. Let i be the payment one gets for the asset. Then the return can be calculated by extending formula (3.1).

$$R = \frac{v_1 - v_0 + i}{v_0} \tag{3.2}$$

This kind of return is also quite straightforward and easy to compute for one period models. [7, p. 862]

3.3 Money-Weighted Return

The money-weighted return, or internal rate of return as it is also called, considers also payments, that have been made inbetween the start and end point of the calculation. Therefore, it is seen as a way to measure return over multiple periods. In general the computation can be done by solving

$$0 = \sum_{t=0}^{T} \frac{c_t}{(1+R)^t} \tag{3.3}$$

for R, where c_t is the cash flow at time t, which can take both negative and positive values. This discount rate is the annual rate of growth that an investment is expected to generate. The higher it is, the more appealing it is to buy said asset. Since it is uniform for various types of investments, it can be used to rank and compare them on a similar basis. The following example demonstrates this return on a two period model.

Example 3.3.1. Let an investor buy a stock for the initial price of 100. After the first period, the asset pays a dividend of 5 and the decision is made to buy another share of said equity. However, the price rose to 103. At the end of the second period both stocks are sold at a price of 104 each.

Time	Outgoing cash flow	Incoming cash flow
0	100	0
1	103	5
2	0	218

Now the average return over this two periods can be calculated using the discounted cash flow and setting the present values of the incoming and outgoing cash:

$$100 + \frac{103}{(1+R)} = \frac{5}{(1+R)} + \frac{218}{(1+R)^2} \tag{3.4}$$

By solving this equation for R the return in this case is approximately 6.57%. Looking at both time steps separately, one can notice, that the bigger impact on the performance comes from the second period. This is due to the fact, that in this time frame the value of the investments is higher and therefore the influence on the overall return is also increased. As of this reason, this measure is called the money-weighted rate of return. [8]

3.4 Time-Weighted Return

Another alternative measure is the time weighted return. In contrast to the money weighted return, this method doesn't depend on the amount invested in the assets. To calculate it, one only needs the total return in each period and then averages over all time periods. So for an investment portfolio over n periods, one needs to take the following formula:

$$R = \frac{1}{n} \sum_{t=1}^{n} R_t \tag{3.5}$$

where the R_t reflects the total return for the period t.

Example 3.4.1. Let's return to the example of the previous section and calculate the time weighted return. We start by looking only at the first period. For the return in that time we get

$$\frac{103 + 5 - 100}{100} = \frac{8}{100} = 8\% \tag{3.6}$$

since we have a starting price of 100 and an end value of 103. By considering also the dividend of 5 we reach a return of 8%. Whereas, in the second period we start with a price of 103 and reach at the end 104 for our two stocks. Therefore, by considering also the total dividend of 10, we end up with the following return.

$$\frac{208 + 10 - 206}{206} = \frac{12}{206} = 5.83\% \tag{3.7}$$

With the last step, taking the average, the calculation comes to a return of 6.91%.

Compared to the money-weighted return this measure is higher due to the fact, that the performance in the first period was better than in the second. Since there was more money invested in the second period, it has more impact in the first calculation.

Due to this fact, the former is also a measure of the investment decision of the manager. It can show if the timing of the purchase was rather good or bad, whereas the time-weighted return doesn't take this into account and only measures the performance of the asset itself. For this reason, it is more commonly used in fields where the portfolio manager doesn't have all the control over the cash flows like in pension funds. In that case, the payments are fixed, but can vary for many reasons, that are out of control of the fund manager. [7, 682 f.]

3.5 Arithmetic and Geometric Average

To round up this section about return measures, we have a look at another approach to the calculation of the time-weighted average. We introduced it as an arithmetic average, while we could also take the geometric version. To calculate it, one again needs the returns over each period of time R_t and plug them into the following formula

$$(1 + R_G)^n = \prod_{t_1}^{n} 1 + R_t \tag{3.8}$$

$$\Leftrightarrow R_G = (\prod_{t_1}^{n} 1 + R_t)^{\frac{1}{n}} - 1 \tag{3.9}$$

This method is also quite interesting since it delivers the constant rate of return one needs to earn yearly to match the actual performance over the investment period.

Example 3.5.1. Let's have another easy example to demonstrate this. A stock with initial price 100 doubles its value and afterwards loses half of it again, so at the end of two periods the price is again at 100. The former arithmetic average delivers a result of $(100\% - 50\%)/2 = 25\%$. However, in reality the value of the asset at the start and the end is the same. The geometric average shows this with a simple computation:

$$1 + R = \sqrt{(1 + 1)(1 - 0.5)} \tag{3.10}$$

$$\Leftrightarrow \quad R = 0 \tag{3.11}$$

As a general property, the geometric average is always lower than the arithmetic one since the bad returns here have a higher impact on the overall performance. The difference between the two can also be seen as a kind of measure of standard deviation. For an asset with a very low variance in its returns, both methods return a similar result, whereas for an investment with very high volatility the difference between the two increases greatly. However, an important property, that distinguishes the two of them, is that the arithmetic average is an unbiased estimator for future performance. On the other side, the geometric average is a downward-biased estimator for the expected return in the future.

To sum it up, the geometric average includes the effect of compounding growth over time. This is important when measuring investment and portfolio performance. [7, 863 f.]

3.6 Expected Return

The question is now what to use for the expected return of an asset in the future. Since this theses aims to combine different asset classes in the optimization process, the answer is not so simple.

For fixed-income investments the historic price is not an appropriated estimator for the expected return since prices and yields are inverse. That means that drops in prices lead to increasing yields and the other way around. Due to this, the yield to maturity (YTM) is a better fitting measure to use. It represents the total rate of return of a bond, that is hold until maturity and fulfills the interest payment as well as the redemption price. In other words, it is the internal rate of return of a bond until maturity transformed into an annual rate. This measure takes into account the current market price P, the coupon payments C, the time to maturity t and the nominal value of the investment N. With this informations the formula to calculate it, is the following

$$YTM = \frac{C + \frac{N-P}{t}}{\frac{N-P}{2}}$$

The big advantage of using this key measure is, that it increases the comparability of bonds with different coupon rates or maturity dates dramatically. Therefore, investors get a better understanding of the investments. [9]

The second type of assets in the construction of an optimal portfolio in this thesis are equities. In this asset class are a few different estimators that can be used. While the mean of the past log returns of an equity can be used for such purposes, a more sophisticated indicator would be the golden cross GC. It is widely known as a signal for an upcoming bull market, i.e. a rising price for the equity. The idea of it is to compare a moving average over the most recent past to the same measure over a bigger time period.

$$SMA = \frac{1}{f_1} \sum_{i=0}^{f_1 - 1} P_t - i$$

$$LMA = \frac{1}{f_2} \sum_{i=0}^{f_2 - 1} P_t - i$$

$$GC = \frac{SMA}{LMA}$$

In this formula SMA stands for the short-term moving average over the last $f_1 \in \mathbb{N}$ timesteps and LMA is the long-term moving average over the last $f_2 \in \mathbb{N}$ ones. The quotient of the two then is the golden cross signal. The investor can then define his expectation of the future return of the asset, for instance the historic mean and multiply it by this factor to include the market trend into the it. [10]

Alternatively for the expected return of equities, one can also use the earnings yield. This key figure measures the earnings per share in the most recent period in relation to the current market price per share. If it is growing, it means the equity generates more income compared to its cost, while small values for the ratio mean that the equity is overvalued. Consequently, this return metric does give a valuable insight on the profitability or the return on an equity investment and can be used to estimate the expected return. [11]

4 Risk Measures

The second side of the coin of portfolio optimization is to evaluate the uncertainity that comes with the various investment. As investors navigate through different asset classes, they encounter numerous risks, from market volatility to economic issues. It is essential to recognize, understand and measure these risk factors in order to construct a portfolio that can withstand different market conditions.

In this chapter, we are going to try to get a comprehensive understanding of risk measures and how they work, from standard deviation and variance to more advanced methods like Value at Risk and Entropy.

4.1 Standard Deviation

The most straightforward form of risk measurement is the standard deviation. In general, it is the spread of data from its expected value. In the context of assets, it is used to measure the historical volatility of an investment in relation to its annual rate of return. In other words, an asset with a high standard deviation varies a lot in its value gains and losses and is therefore considered riskier. As one might already know, the computation of the standard deviation σ is quite commonly known as the root of the variance $\mathbb{V}$

$$\mathbb{V}(X) = \sigma^2 = \mathbb{E}[(X - \mu)^2] \tag{4.1}$$

where μ is the expected value of the return and the random variable X describing the return of the asset.

If one wants to get to concentrate on the riskier side of the deviation, it is also possible to exclude the gains and only consider the losses by using the negative part of the expectation. The argument here is that all deviations above the mean are desirable and therefore not risky.

$$\mathbb{V}_{semi}(X) = \mathbb{E}[((X - \mu)_-)^2] \tag{4.2}$$

This is called the semivariance. However, it is much more difficult to use, when calculating it for whole portfolios and not for single investments. Furthermore, empirical data shows that the returns of most stocks are reasonably symmetrical, which causes semivariance to be proportional to the normal variance and standard deviation. For this reason, it is not very widespread. [12][p.49 ff.]

4.2 Value at Risk

As it is already known, the return of an asset or even a whole portfolio is expressed as a random variable X that depends on a single or multiple risk factors, i.e. prices, exchange rates or interest rates. The expected value and standard deviation can be estimated using historical data. From this definition, one can also derive the losses L over a period of time as a random variable:

$$L_{t,t+\delta} = X_t - X_{t+\delta} \tag{4.3}$$

Starting from the underlying distribution of X, the distribution of L can also be determined as dependent on the information available at time t.

In practice, one of the most common risk measures is the Value-at-Risk(VaR) of the loss. For a chosen confidence level α, the VaR is defined by the smallest value l for which the probability of a loss L is lower than $1-\alpha$ and the loss itself is bigger than l. In other words, it is the α-quantile of the loss.

$$VaR_\alpha = \inf\{l \in \mathbb{R} : \mathbb{P}(L > l) \leq 1-\alpha\} \tag{4.4}$$
$$= \inf\{l \in \mathbb{R} : 1 - F_L(l) \leq 1-\alpha\} \tag{4.5}$$
$$= \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\} \tag{4.6}$$

Example 4.2.1. A one-period 95% VaR of $l = 100000$ means that in 95 cases out of 100 the loss over the time period will not exceed this level l.

Although, this measure is in reality widely used since it has a lot of useful characteristics like monotonicity, homogeneity and translation invariance, it does come with some limitations and downsides. First of all, it assumes the distribution of returns is known and constant, which is in reality not applicable, because in most cases there are widely spread portfolios with all kind of assets. It can be quite difficult, if not impossible, to determine the mixed distribution. Moreover, it does not give any indications on how big the loss is going to be, if it exceeds the given level, which might be very bad in case of heavy-tailed distributions. Also it does not consider the timing of losses, which can be crucial in some situations. [13][35]

4.3 Conditional Value at Risk

The conditional Value-at-Risk (CVaR) is an extension of the normal VaR to assess the risk in extreme scenarios. It measures the size of the average losses that exceed the VaR for a given confidence level. Formally spoken, it can be expressed as a conditional expectation.

$$CVaR_\alpha = \mathbb{E}[L | L > VaR_\alpha] \tag{4.7}$$

It is very useful for capturing the tail risk or, in other words, it measures the risk of extreme losses, which the standard VaR might underestimate. Furthermore, it is a coherent risk measure, i.e. it satisfies subadditivity, translation invariance, positive homogeneity and monotonicity.
For continuous distributions, the CVaR takes on the form

$$CVaR_\alpha = \frac{1}{1-\alpha} \int_\alpha^1 VaR_u du \tag{4.8}$$

while for discrete versions the formula is the following

$$CVaR_\alpha = \frac{1}{1-\alpha} \sum_{j:L_j \geq VaR_\alpha} L_j p_j. \tag{4.9}$$

Example 4.3.1. Given an empirical distribution with a sample size of 1000, the CVaR for a confidence level of 99% is easily computed, by taking the 10 biggest losses and calculating the average of those.
In practice, this measure is a very common and important one since it gives investors a better understanding of risks that go beyond the VaR level. It provides a more comprehensive view of extreme risk, so to speak. [13][36]

4.4 Beta

Another approach to measure how risky an asset is, is to look at its Beta. This concept takes into account how volatile the investment is compared to the market. In this context, market is meant to be an index or some sort of benchmark, for instance the S&P 500 could be an appropriate choice for a stock.

To calculate the Beta, one only needs to know two things, the variance of the market itself $\mathbb{V}(X_M)$ and the covariance between it and the chosen asset $\text{Cov}(X_M, X)$, where X_M represents the returns of the market as described above and X the ones for said asset. The former variance is also referred to as the systematic market risk. The ratio of the two of them is the Beta.

$$\beta = \frac{\text{Cov}(X_M, X)}{\mathbb{V}(X_M)} \tag{4.10}$$

An asset with $\beta = 1$ moves pretty much along with the market as it has the same variance. For values above 1 an asset would be more volatile than the benchmark, which means it comes with more risk, but also with more possible reward in terms of returns given a rising market. Conversely, investments with low Betas pose less risk and also lower potential gains. This method of risk measurement could also return negative values. This would mean that the asset moves against the swing of the market or in other words is negatively correlated to market movements.

Due to its properties, conservative investors are more likely to look for the stocks or bonds with low Betas, while day-traders, who look for fast and quick gains, are more interested in high Beta assets. This measure is primarily useful when considering the capital asset pricing method (CAPM), which is introduced later on in chapter 5.2.

Although Beta does have some very useful properties and returns interesting information on an investment, it also does come with some disadvantages. Since it is calculated by using historical data points, Beta is not appropriate to determine the future movements of the stock. Due to this and the fact, that the volatility of an asset can change yearly, making it an unstable measure, it is not useful when considering long-term investments. [14]

4.5 Alpha

To conclude this chapter about risk measures, we have a look at the Alpha of an investment or fund. The Alpha-measure is often used when describing the strategy's ability to outperform the market itself over a certain period of past time. Therefore it is also often called the "excess return" or "abnormal rate of return" in relation to a benchmark. In other words, it is not a risk measure per se, but more like a risk-adjusted performance measure of the portfolio manager, since it is a result of active investing. An Alpha value of zero would indicate, that the manager did not add or lose any additional value in the specified time period and is tracking the market perfectly. [15]

The reason this measure is included in this chapter, is that it is based on the Beta, which was described above and is often used in combination with it. To calculate it, one needs to start off with determining the excess return the benchmark achieved over a risk-free investment. In other words, this is simply the difference between the return of the former minus the return of the latter. With this information, the next step is to compute the expected return of the market by multiplying the excess with the Beta - that got introduced in the previous section - of the portfolio or investment at hand and adding it to the risk-free return.

$$R_M = R_{RFR} + \beta \times (R_{BM} - R_{RFR}) \tag{4.11}$$

In this formula, the R in general stands for the expected return. The indices state which of the components they represent, namely either the risk-free rate (RFR), the benchmark (BM) or the market (M). The second term on the right, $\beta \times (R_{BM} - R_{RFR})$, is also called the risk premium since it illustrates the additional return, that is required to justify the risk taken by investing in that portfolio.

Now, all that is left to do is to calculate the difference between the actual return of the portfolio R_{PF} and the expected return of the market R_M.

$$\alpha = R_{Pf} - R_M \tag{4.12}$$

As we already know from the previous chapter, a high Beta value indicates a more volatile asset or fund. Such a portfolio does have to perform a whole lot better to reach a positive alpha, since the expected return is directly correlated to the Beta and therefore leads to a higher risk-adjustment. Two funds could come with the same return, but with quite different alphas due to different risk levels.

As all other risk, return or performance measures, this one is also not perfect. First and foremost, the same limitation of Beta also get inherited by the calculation of Alpha. It is a measure of the past success and thus not too reliable as a future predictor. It can be influenced by luck, a few successful bets of the portfolio manager or also a combination of both with a good market timing.

Moreover, the calculation of Alpha is also quite sensitive to the assumptions made as the choice of the risk-free rate or the benchmark. A badly chosen benchmark, for instance one that has a low correlation to the fund, can return a misleading value for Alpha. Additionally, the choice of the time period is very important. For too short-term periods, the Alpha can include some noise and randomness and may not reflect the actual performance of the fund or the skill of its manager. [16]

5 Portfolio Models

In this chapter, we dive into the world of different theories on how to get a optimal portfolio. Since there are countless different approaches, we concentrate on the most common and important ones.

5.1 Markowitz Model

To start things off, we have to talk about the father of the Modern Portfolio Theory as we know it. In 1952 Harry Markowitz introduced his groundbreaking article "Portfolio Selection", in which he shared his insights on risk, return and the benefits of diversification. Those thoughts have altered the way investors and researchers think about said aspects, shaping the field of portfolio management until today. [1]

5.1.1 Introduction

In his article, Markowitz addressed the universal problem of investors, namely to maximize the return of their portfolio all while minimizing their risk. He defined risk as the variability of the return and stated that a key aspect is the trade-off between that and the return itself. Markowitz was already aware that assets with greater returns also pose higher risk levels and investors have to take that into account when deciding on investments.

His revolutionary thougth was that the risk/return-profile of the separate assets should not be examined for each on its own, but alltogether as a whole portfolio. Markowitz claimed those portfolios to be efficient if they either maximize the return for a given level of risk or vice versa minimize the risk for a given level of return. Although, both approaches are equivalent to each other, the type of optimization problem is a different one. The first one turns out to be linear optimization with quadratic constraints whereas the second one is a quadratic one with linear constraints. [13][p.43 f.]

5.1.2 Mean-Variance Analysis

To deal with this problem, he introduced the mean-variance analysis. Markowitz assumed that he has N assets to choose from with every single one being divisible indefinitely and the returns from those to be jointly normally distributed. The concept relies on three key figures to assess the risk and return, the expected return $\mathbb{E}(R_i)$, the variances $\mathbb{V}(R_i)$ and the covariances $\text{Cov}(R_i, R_j)$ of every asset i. Based on those the return and the variance of the whole portfolio can be calculated by applying the following formulas:

$$\mathbb{E}[R] = \sum_i \omega_i \mathbb{E}[R_i] \tag{5.1}$$

$$\mathbb{V}(R) = \sum_i \sum_j \omega_i \omega_j \text{Cov}(R_i, R_j) \tag{5.2}$$

In this equations, ω_i stands for the weight of the asset i and the covariance being calculated by the well known formula

$$\text{Cov}(R_i, R_j) = \rho_{ij}\sigma_i\sigma_j \tag{5.3}$$

with ρ_{ij} being the correlation coefficient of assets i and j. In the financial world, covariance measures the direction in which the returns of two assets move. With positive values, both tend to go in the same direction, while negative values imply that two assets trend into opposite ones.

With this formulae, one can specify the optimization problem in some different ways depending on what exactly to achieve. The first possibility is to maximize the return of the portfolio given a specific level σ^2 of variance.

$$max\ \mathbb{E}[R] \tag{5.4}$$

$$s.t.\quad \mathbb{V}(R) \leq \sigma^2 \tag{5.5}$$

$$and\quad \sum_i \omega_i = 1 \quad with\ 0 \leq p_i < 1,\ i = 1, 2, ..., n. \tag{5.6}$$

The goal is here to reach as much return as possible while maintaining at least an upper border for the variance. Without it, the optimization would invest the whole capital in the asset with the highest return expectation, making it not applicable. Moreover, this definition of the optimization problem is more theoretical and not used very often in reality.

The other direction is the more common approach of the Markowitz model. This way the variance is minimized with the optional constraint of a minimum return. The resulting optimization problem is a quadratic one with linear constraints.

$$min\ \mathbb{V}(R) \tag{5.7}$$

$$s.t.\quad \mathbb{E}[R] \geq \mu \tag{5.8}$$

$$and\quad \sum_i \omega_i = 1 \quad with\ 0 \leq p_i < 1,\ i = 1, 2, ..., n. \tag{5.9}$$

Theorem 5.1.1. *The solution to the optimization problem in (5.7-5.9) has the form*

$$\omega = (1 - \alpha)\omega_{min_var} + \alpha\omega_{mk}$$

,where ω_{minvar} are the weights of the global minimum variance portfolio, ω_{mk} are the market weights with the form $\omega_m k = \frac{\Sigma^{-1}\mathbb{E}[R]}{\mathbf{1}^T\Sigma^{-1}\mathbb{E}[R]}$ and

$$\alpha = \frac{\mu(\mathbb{E}[R]^T\Sigma^{-1}\mathbb{E}[R])(\mathbf{1}^T\Sigma^{-1}\mathbf{1}) - (\mathbb{E}[R]^T\Sigma^{-1}\mathbf{1})^2}{\delta}.$$

Proof. The objective function of the problem is to minimize the variance σ^2. Since Σ is positive definite, the variance term is strictly convex. To solve such a constrained optimization problem, one can use the method of Lagrange multipliers. The Langrangian function is defined as

$$L(\omega, \lambda_1, \lambda_2) = \omega^T \Sigma \omega + \lambda_1(\mu - \mathbb{E}[R])^T \omega + \lambda_2(1 - \mathbf{1}^T \omega),$$

where the first term corresponds to our objective function, i.e. the variance. The second term is the constraint of a minimum return and the last illustrates the full investment constraint, i.e. that the sum of the weights has to be equal to one. Then, we can extract the Karush-Khun-Tucker(KKT) conditions.

$$0 = \Sigma\omega - \lambda_1 \mathbb{E}(R) - \lambda_2 \mathbf{1} \tag{5.10}$$

$$\mu \leq \mathbb{E}[R]^T \omega, \quad \mathbf{1}^T \omega = 1, \quad 0 \leq \lambda_1 \tag{5.11}$$

$$\lambda_1(\mathbb{E}[R]^T \omega - \mu) = 0 \tag{5.12}$$

Now, let $\hat{\omega}$ be the solution to the problem and we can see that we have two cases:

- $\mu < \mathbb{E}[R]^T \hat{\omega}$: This would imply that λ_1 is equal to 0 due to the third KKT condition leaving only two equations:

  $$0 = \Sigma\hat{\omega} - \lambda_2 \mathbf{1}$$
  $$\mathbf{1}^T \omega = 1$$

  Then, we can multiply the first by the inverse of the covariance matrix Σ and with the unity vector $\mathbf{1}$ to get $\mathbf{1}^T \hat{\omega} = \lambda_2(\mathbf{1}^T \Sigma^{-1} \mathbf{1})$. Since the left side is equal to 1 due to the second KKT condition we end up with $\lambda_2 = (\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1}$ and can plug this into the previous equation to get

  $$\hat{\omega} = (\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1} \Sigma^{-1} \mathbf{1}.$$

  Therefore, all KKT conditions are true for this $\hat{\omega}$ making it in fact a solution to the optimization problem. It is important to note, that this solution yields the smallest possible variance over all portfolios.

- $\mu = \mathbb{E}[R]^T \hat{\omega}$ In this case, we start by dividing the first KKT condition by Σ.

  $$\hat{\omega} = \lambda_1 \Sigma^{-1} \mathbb{E}[R] + \lambda_2 \Sigma^{-1} \mathbf{1} \tag{5.13}$$

  This expression then can be used for the second line of the KKT conditions.

  $$\mu = \lambda_1 \mathbb{E}[R]^T \Sigma^{-1} \mathbb{E}[R] + \lambda_2 \mathbb{E}[R]^T \Sigma^{-1} \mathbf{1}$$
  $$\mathbf{1} = \lambda_1 \mathbb{E}[R]^T \Sigma^{-1} \mathbf{1} + \lambda_2 \mathbf{1}^T \Sigma^{-1} \mathbf{1}$$

  Extracting the left sides and also λ_1 and λ_2 as vectors, these equations can be written in 2×2 matrix equation with the matrix having the form

  $$T = \begin{bmatrix} \mathbb{E}[R]^T \Sigma^{-1} \mathbb{E}[R] & \mathbb{E}[R]^T \Sigma^{-1} \mathbf{1} \\ \mathbb{E}[R]^T \Sigma^{-1} \mathbf{1} & \mathbf{1}^T \Sigma^{-1} \mathbf{1} \end{bmatrix}$$
  $$= \left[\mathbb{E}[R] \mathbf{1}\right]^T \Sigma^{-1} \left[\mathbb{E}[R] \mathbf{1}\right]$$

This matrix is always is always positive semi-definite since Σ is also positive definite and the following holds, if $\mathbb{E}[R]$ and $\mathbf{1}$ are linear independent:

$$0 < \delta = (\mathbb{E}[R]^T\Sigma^{-1}\mathbb{E}[R])(\mathbf{1}^T\Sigma^{-1}\mathbf{1}) - (\mathbb{E}[R]^T\Sigma^{-1}\mathbf{1})^2$$

To see that this is true, one has to look at $\delta = 0$. Then the expected return has to be equal to $\tau\mathbf{1}$ for some $\tau \in \mathbb{R}$. However, this would mean for $\mu/\tau \neq 1$, that the problem is necessarily infeasible. On the other hand for $\mu/\tau = 1$, there would only be the minimum variance solution since every portfolio delivers the same expected return.

If $\delta > 0$, the system described above can be solved and we can express λ_1 and λ_2 in the following way:

$$\begin{aligned} \lambda_1 &= \mathbf{1}^T\nu \\ \lambda_2 &= -\mathbb{E}[R]^T\nu \\ \text{with} \quad \nu &= \delta^{-1}\Sigma^{-1}(\mu\mathbf{1} - \mathbb{E}[R] \end{aligned}$$

As a consequence, we can resubstitute this into 5.13 and get the optimal solution for our problem:

$$\begin{aligned} \hat{\omega} &= \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}} + \alpha\left[\frac{\Sigma^{-1}\mathbb{E}[R]}{\mathbf{1}^T\Sigma^{-1}\mathbb{E}[R]} - \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}\right] \\ &= (1-\alpha)\frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}} + \alpha\frac{\Sigma^{-1}\mathbb{E}[R]}{\mathbf{1}^T\Sigma^{-1}\mathbb{E}[R]} \\ &= (1-\alpha)\omega_{min_var} + \alpha\omega_{mk}. \end{aligned}$$

□

5.1.3 Diversification

The key aspect to this whole model is the importance of diversification. Markowitz states that investors can eliminate the risk of a portfolio up to a certain point without losing out on the return by mixing assets with non-perfectly correlated returns. This is due to the fact, that the portfolio risk does not only rely on that of each single component, but also on the interaction between them. In other words, the correlation between assets needs to be estimated as well to do mean-variance optimization according to Markowitz.

Furthermore, Markowitz also gives two logical restrictions to the weight of the assets. The first is that the sum of them should be one while the second states that no negative weights are allowed. The former just adds a little bit of normalization to the weights. The latter however, prevents the model from returning infinite weights since it could just short sell an infinite amount of undesirable assets and buy other ones with that. This could also be weakened by restricting short selling, i.e. negative weights, only up to a certain level. [1]

5.1.4 The Efficient Frontier

Last but not least, Markowitz introduced visualizations of all possible portfolios in terms of risk and return. Going along the border of the set of all possible portfolios, one can find all of the efficient ones, i.e. the portfolios with maximum return for a given level of risk or vice versa minimum risk for a given return level.

Theorem 5.1.2 (Caratheodory). *Let S be a convex subset of $\mathbb{R}^n$. Every element $s \in S$ can be described as a convex combination of $n+1$ elements of S,*

$$s = \sum_{i=1}^{n+1} \lambda_i s_i \quad \textit{with } s_i \in \Omega, \ \sum_{i=1}^{n+1} \lambda_i = 1, \ \lambda_i \geq 0. \tag{5.14}$$

Proof. Suppose s is a convex combination of the form

$$s = \sum_{i=1}^{k} \lambda_i s_i \ \text{ with } s_i \in \Omega, \ \sum_{i=1}^{k} \lambda_i = 1, \ \lambda_i \geq 0$$

with $k > n+1$. since there are more than $n+1$ elements, they are linearly dependent. Therefore, we can look at the homogeneous linear system of equations

$$\alpha_1 s_1 + \alpha_2 s_2 + ... + \alpha_k s_k = 0$$
$$\alpha_1 + \alpha_2 + ... + \alpha_k = 0.$$

Due to the fact that $s_i \in \mathbb{R}^n$, this system has $n+1$ equations and therefore less than the number of indeterminates. This ensures that there is a non-trivial solution $(\alpha_1, \alpha_2, ...\alpha_k)$.

As the λ_i sum up to 0, but not all are equal to 0, there exists at least one i with $\lambda_i > 0$. Let

$$\tau = min\{\frac{\lambda_i}{\alpha_i} | \alpha_i > 0\}$$

and

$$\beta_i = \lambda_i - \tau\alpha_i \forall i = 1, ..., k.$$

Then, the following holds

$$\beta_i \geq \lambda_i - \frac{\lambda_i}{\alpha_i} = 0$$
$$\sum_{i=1}^{k} \beta_i = \sum_{i=1}^{k} \lambda_i - \tau \sum_{i=1}^{k} \alpha_i = 1.$$

Now, we can rewrite the convex combination from the start.

$$\beta_1 s_1 + ...\beta_k s_k = \lambda_1 s_1 + ...\lambda_m s_m - \tau(\alpha_1 s_1 + ...\alpha_m s_m) = x$$

This results in our goal. When there is an index j with $\tau = \frac{\lambda_j}{\alpha_j}$, then $\beta_j = 0$ due to the construction and we can cut off s_j. □

This theorem can be applied to the set of all portfolios, since this is also a convex subset of the $\mathbb{R}^2$, characterized by the return and the risk. With this, we can derive also a theorem for the upper border of this convex set.

Theorem 5.1.3 (Two-fund theorem). *Given two portfolios on the efficient frontier, one can express any other efficient portfolio as a linear combination of the first two.*

Proof. Let's assume, there are two portfolios on the efficient frontier characterized by the vectors of their asset weights ω_A and ω_B. Their according expected returns are then given by μ_A and μ_B while their variances are σ_A^2 and σ_B^2. The corresponding covariance is also known by σ_{AB}.
Now consider a new portfolio as a linear combination of the two of them with some scalar α.

$$\omega_C = \alpha\omega_A + (1-\alpha)\omega_B \tag{5.15}$$
$$\tag{5.16}$$

The expected return can thus be calculated with the following formulas.

$$\begin{aligned}
\mu_C &= \alpha\mu_A + (1-\alpha)\mu_B && (5.17)\\
\sigma_C^2 &= (\alpha\omega_A + (1-\alpha)\omega_B)^T\Sigma(\alpha\omega_A + (1-\alpha)\omega_B) && (5.18)\\
&= \alpha^2\omega_A^T\Sigma\omega_A + 2\alpha(1-\alpha)\omega_A^T\Sigma\omega_B + (1-\alpha)^2\omega_B^T\Sigma\omega_B && (5.19)\\
&= \alpha^2\sigma_A^2 + 2\alpha(1-\alpha)\sigma_{AB} + (1-\alpha)^2\sigma_B^2 && (5.20)
\end{aligned}$$

Due to the convexity of the efficient frontier, the new portfolio ω_C either lies on this line or is within the boundary of it. Additionally, the two initial portfolios are both efficient, hence they deliver the best possible expected return for their levels of risk. By combining them, the resulting portfolio also inherits this property. Therefore, it maintains this optimal risk-return trade-off and thus lies on the efficient frontier as well. □

As a result, any efficient portfolio on this frontier can be derived by a linear combination of the minimum variance portfolio and another efficient portfolio on it. The covariance between the derived portfolio and the chosen efficient one is then equal to the variance of the minimum variance portfolio.

From looking at figure 5.1, one can see that there indeed exists an unique global minimum for the risk. Extending this graph to the right, it can also be seen, that this set of portfolios is closed and there also exists a global maximum return portfolio. [13][p.44 f.]

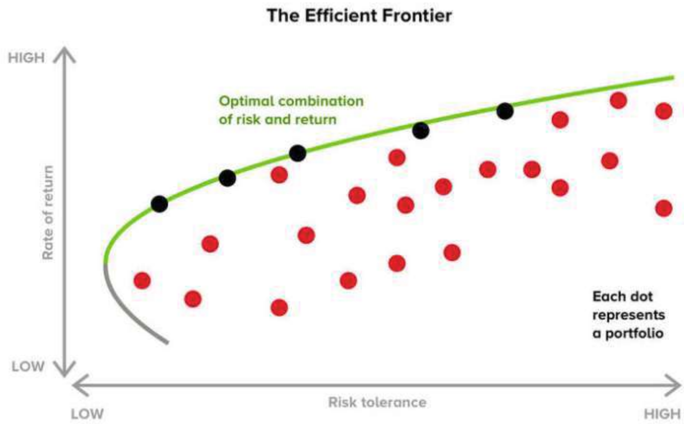

Figure 5.1: Example of an efficient frontier[17]

5.1.5 Conclusion and Problems

To sum up, Markowitz's model provides portfolio managers with a systematic and balanced approach by considering the trade-off between risk and return. The concepts of mean-variance analysis, efficient frontier, diversification and covariances help investors a lot with constructing a balanced and robust portfolio. Its impact cannot be understated since so many more portfolio models and risk management strategies are based on it and to this day it is still one of the most important models in existence.

On the other hand, there are some problems that can be encountered in the real world when using the Markowitz model. In practice, the mean-variance optimization can lead to highly concentrated allocations in very few assets, especially when the goal is to reach as much return as possible. In such extreme cases, the solution tends to lose the desirable effect of the diversification and is to invest in a very limited amount of possibly extremely risky assets. Further on, investors want to include their own views on the market in the model by adjusting some inputs to the calculation of the covariance matrix or choosing a different way to calculate it and thus can disturb the dependence structure. This can lead to irritating and highly volatile optimization results. Even slight changes have the possibility to do a lot of harm and alter everything in unpredictable ways.

Another difficulty of this approach is the assumption of normality among the asset returns. If they appear to be non-normal or asymmetric, for instance it the returns follow a log-normal or Student's t-distribution, the optimization can result in surprising and unexpected portfolios. In fact, most returns are not symmetrically in practice, which makes the variance a poor risk measure and leads to losing upside volatility in returns. In addition to all of this, portfolio optimization models have the claim to be forward-looking as they want to project the best portfolio for the future. However, the estimators for the covariance and the return rely on historical data and therefore can often turn out to be not that useful. [3]

To tackle all this issues, we examine some different models in the upcoming sections and then try to see which one would be useful for deciding on future investments.

5.2 The Capital Asset Pricing Model

The next very important model is the Capital Asset Pricing Model (CAPM), which is an extansion of the Markowitz model. It was introduced in 1964 by an American economist called William F. Sharpe, who also won the Nobel Prize in Economic Sciences for his work on the CAPM together with Harry Markowitz and Merton Miller. The key aspect here is the calculation of the expected return of a risky asset by implying a linear relationship between it and the expected return of the market. This is done by adding the parameter β to the calculation, but first let's talk about the motivation.

5.2.1 Assumptions

Before we start off with the motivation and idea behind the CAPM, we need to clarify some assumptions made by it. First off, the model makes a few such premises about the market circumstances itself. It assumes that there are neither transaction costs nor taxes on the gains from assets. Both of those would add a lot of unwanted complexity to the model, since the goal of this paper is not to address problems in relation to taxes or costs of buying assets. The next premise is that all assets are infinitely divisible, i.e. that an investor can invest any amount of money in a stock and does not have to consider the stock prices. Furthermore, all market prices are independent from single market participants. In other words, a single investor cannot influence the price by selling or buying for instance a certain stock. The prices are just regulated by the overall demand of all investors together. Additionally, investors are allowed to make unlimited short sales and there is also the possibility of unlimited lending and borrowing at a riskless rate. Finally, the CAPM assumes that all assets, including human capital, are marketable or in other words can be bought and sold.

On the other hand, there are also a few assumptions made about the market participants themselves. It is assumed, that all investors make their respective decisions solely based on the expected return and standard deviation of the returns of the assets. Moreover, all market participants also share the same believe on the distributions of those key figures and they also define the relevant period of the investment exactly the same way.

To sum up, there are quite a few assumptions to the CAPM that are quite obviously not applicable in the real world. Nonetheless, the CAPM still delivers a quite good view on the risk and return itself but more on that in the following. [12][p.294 f.]

5.2.2 Motivation

Let's assume that the weights of a portfolio with n assets have the form $w_i = \frac{1}{n}, 1 \leq i \leq n$. The goal is to calculate the variance of the whole portfolio. To do this, we need to have a look at the variances and covariances of the assets. Suppose the former are uniformly bounded by $\sigma^2 \leq L$. Then we can plug into our formula for the portfolio risk.

$$\sigma^2 = \sum_{i,j=1}^{n} \omega_i \omega_j \sigma_{ij} = \sum_{i=1}^{n} \omega_i^2 \sigma_i^2 + \sum_{i \neq j} \omega_i \omega_j \sigma_{ij} \leq n \frac{1}{n^2} L + \frac{1}{n^2} \sum_{i \neq j} \sigma_{ij} \tag{5.21}$$

Looking at the covariance matrix, the elements in the diagonal are already bounded since those are the variances themselves. A similar assumption is now made for the off-diagonal elements, namely that they are also uniformly bounded by a constant $c \geq |\sigma_{ij}|, i \neq j$. This leads to

$$\sigma^2 \leq \frac{L}{n} + \frac{1}{n^2}n(n-1)c \tag{5.22}$$

Finally, let n go to ∞ and one can see that the upper bound for the risk of a portfolio converges to c. This leads to the conclusion, that the variances of single assets become irrelevant for a very wide spread portfolio since the first term in the inequality is the upper bound of those and converges to zero for high n.

This proof motivates a differentiation between two different types of risk. The first one is the diversifiable or specific risk. That risk can be reduced to zero by adding assets and expanding the portfolio. The other one is called the undiversifiable, systematic or market risk. This one describes the interaction between the assets and their movements together. All in all, this idea of separating the risk implies that instead of minimizing the overall variance, one should concentrate on the undiversifiable risk, which in turn depends on the covariances between the assets. The CAPM tries to do exactly that by linking the systematic risk of an asset with its expected return. [18][p.67 ff.]

5.2.3 Formula and Calculation

To start the calculation itself, we first have a look at the computation of the expected return. Suppose that there exists a risk-free asset with a guaranteed return R, which is lower than the expected return of the whole portfolio, and has, as the name says, no risk. According to the CAPM, the expected return μ_i of the i-th asset in the portfolio can be calculated through the formula

$$\mu_i = R + \beta_i(\mu_m - R). \tag{5.23}$$

using the β_i of each asset, which got introduced in section 4.4. This formula might seem familiar since we already introduced it in section 4.5 to calculate the alpha of a portfolio. As mentioned before the beta determines the risk premium an investor has to get to justify the taken risk and as a result also the expected return on a security. This also means that beta quantifies the undiversifiable risk.

With this information, one can proceed with the same steps as in the Markowitz model and calculate the minimum variance portfolio and efficient frontier using this expected returns instead of the mean returns in the previous section. Additionally to these key figures, the CAPM also uses the capital market line, the security market line and the Sharpe ratio as indicators for investment decisions. [18][p.67 ff.]

5.2.4 The Market Portfolio

According to the CAPM, every portfolio manager holds a combination of the risk-free asset and the market portfolio, which is the optimal portfolio on the efficient frontier when one considers the existence of a risk-free asset. Basically, the market portfolio contains all risky assets in the market, weighted according to their market values. It serves as a representation of the overall market. Since the weights are the market values of the assets, the bigger a company and its market capitalization is, the higher is its impact on the performance of the market portfolio.

The goal of the market portfolio is to provide a very well-diversified representation of the entire market by including all possible risky assets. This is done to spread the risk across different assets and reduce the impact of poor performance in a single asset. In reality, such a portfolio is nearly impossible to implement or generate and therefore market indices, such as S&P 500 or MSCI World, are used as market portfolios. The market portfolio and thus the capital market line are a feasible theory only if the risk-free return R is smaller than the expected return of the minimum variance portfolio. Else an investor would have no interest in taking on some risk if he can get more reward when taking no risk. In reality this assumptions hold in nearly all cases, since very secure assets usually do not hold remarkable returns.[18][p.67 ff.]

Now, we do a little excursion into linear regression. Let X and Y be random variables taking real values as well as Z be the n-dimensional vector-valued random variable. the next step is to perform a linear regression of both X and Y on Z. To do this, we look at the corresponding model

$$X = \alpha_x + \beta_x Z + Res_x$$

and minimize the sum of squared residuals

$$Res_x = \min_{\alpha_X, \beta_X} \sum_{i=1}^{n} (X_i - \alpha_X - \beta_X Z_i)^2$$

by using ordinary least squares. This results in the residuals having the form

$$Res_x = X - (\alpha_X + \beta_X Z).$$

The analogue procedure is for the regression of Y on Z. Now, we calculate the correlation between those residuals and end up with the following definition.

Definition 5.2.1 (Partial Correlation Coefficient)**.** The partial correlation between two random variables X and Y given a set of controlling variables Z is defined as the correlation between the residuals of X and Y.

$$\rho_{XY \cdot Z} = \frac{\rho_{XY} - \rho_{XZ}\rho_{YZ}}{\sqrt{(1 - \rho_{XZ}^2)(1 - \rho_{YZ}^2)}} \tag{5.24}$$

The partial correlation coefficient then can be used to calculate the weights of the market portfolio by using the following formula

$$m = \frac{C^{-1}(\mu - R\mathbf{1})}{\mathbf{1}^T C^{-1}(\mu - R\mathbf{1})} \tag{5.25}$$

where C is the covariance matrix of the whole market, μ is the vector of the individual returns and $\mathbf{1}$ is a vector of just ones. Moving forward, one can calculate the expected return μ_m of this portfolio m.

5.2.5 The Capital Market Line

Now, we can use this μ_m to get a more general form of the formula 5.23. With the same assumption about the risk-free rate, for any portfolio w holds

$$\mu_w = R + \beta_w(\mu_m - R). \tag{5.26}$$

By using this expectation and the according standard deviation σ_m, we can draw a line through the points $(0, R)$ and (σ_m, μ_m) or in the form of an equation

$$\mu = R + \frac{\mu_m - R}{\sigma_m}\sigma. \tag{5.27}$$

This graph is called the capital market line (CML). The last term on the right $\frac{\mu_m - R}{\sigma_m}\sigma$ is also called the risk premium.

If all investors use the mean-variance optimization and believe in the same parameters for the model, i.e. the same expected returns on the assets and same covariance matrix. Then all optimal portfolios would lie on the capital market line.

Since all investors share the same beliefs on the expected returns and the covariance matrix, they would all compute the same efficient frontier as well as the same capital market line for everybody. To achieve the individual portfolio ω_{PF} with the desired risk-return profile, each investor would buy a combination of the risk-free asset ω_{RFR} and the risky portfolio, i.e. the market portfolio ω_M. Therefore, the calculation is very straightforward with the parameter α.

$$\omega_{PF} = \alpha\omega_{RFR} + (1 - \alpha)\omega_M \tag{5.28}$$

One can see easily that this linear combination never leaves the capital market line, which is drawn straight from the risk-free asset to the market portfolio.

In other words, every investor should invest in the mix of the risk-free asset and the market portfolio. The only difference between them would be their risk aversion, so how they weight these two ingredients.

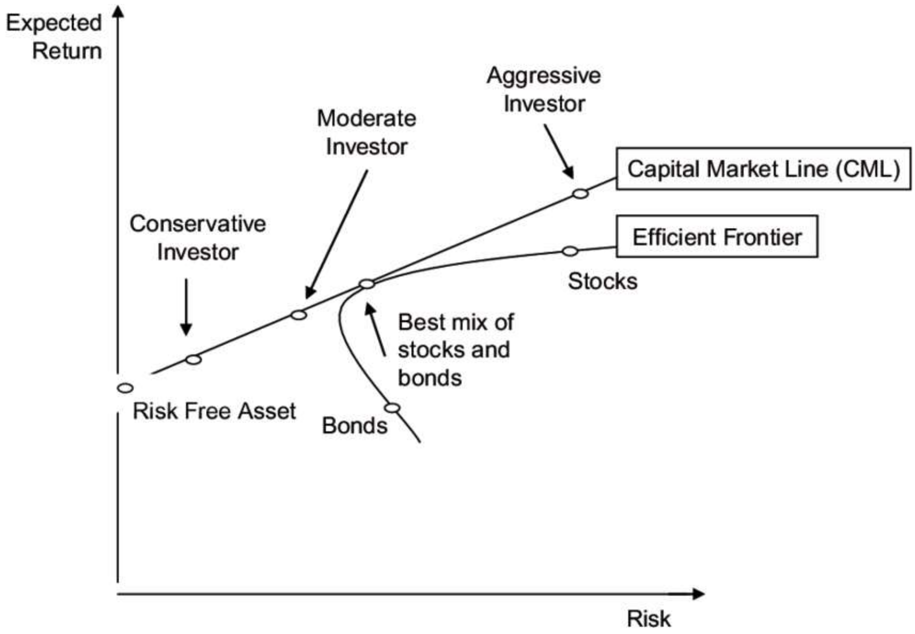

Figure 5.2: Example of the Capital Market Line[19]

As one can see in figure 5.2 the Capital Market Line is tangent to the efficient frontier and the point of tangency is exactly the market portfolio. Depending on the risk appetite of an investor, their portfolio is more to the left, if they are more conservative and risk averse and more to the right, if they are more aggressive in their approach of investing. [18][p.62 ff.]

5.2.6 The Sharpe Ratio

The slope of the capital market line is also called the Sharpe Ratio due to its discoverer, the before mentioned William F. Sharpe. It is a risk-adjusted performance measure since it takes both the return and the risk of an investment or portfolio into account. This key figure helps an investor to evaluate if the additional risk taken is justified by the additional return of the security and is a way to compare investments on a risk-adjusted basis. To put it more formally, it measures the return, that exceeds the performance of the benchmark or risk-free rate, per unit of risk taken. Therefore, a high Sharpe Ratio, usually above 1, is desirable. The formula is implicitly already described in equation 5.27

$$SR_m = \frac{\mu_m - R}{\sigma_m}. \tag{5.29}$$

for the benchmark, i.e. the market portfolio. The general form for a portfolio w is

$$SR_w = \frac{\mu_w - R}{\sigma_w}. \tag{5.30}$$

The optimal Sharpe Ratio is the one for the market portfolio. This portfolio gathers the most additional return over the risk free rate per unit of risk that the investor takes on.

Example 5.2.2. Let the risk-free rate R be 2%. Assume the expected return of the market portfolio μ_m is 5% and its standard deviation σ_m is 15%. Then the Sharpe Ratio can be calculated.

$$SR_m = \frac{0.05 - 0.02}{0.15} = 0.02 \tag{5.31}$$

Therefore, the Sharpe Ratio of the market portfolio is 2% and at the same time optimal. The investor with this portfolio strategy gathers 2% more return than the risk-free rate for every percent of its standard deviation. Every other portfolio is below this threshold and delivers less return per unit of risk.

This key figure has in practice a quite important role. It is used by many investors, who seek to assess the risk-adjusted performance and compare investments. To sum it up, the Sharpe Ratio is a very powerful tool to make investment decisions based on the risk-return trade-off of assets. However, one should also keep in mind that the standard deviation does not capture all aspects of risk, since financial assets are often not normally distributed. Due to this reason, there are also strategies with high Sharpe Ratios that lead to losses. [20]

5.2.7 The Security Market Line

The next indicator derived from the CAPM is the security market line (SML). This graph is quite similar to the capital market line but instead of the (σ, μ)-plane we have a look at the (β, μ)-world. Using formula 5.23 one can draw a straight line there and see that the intercept with the y-axis is again the risk-free rate R and the slope is $(\mu_m - R)$.

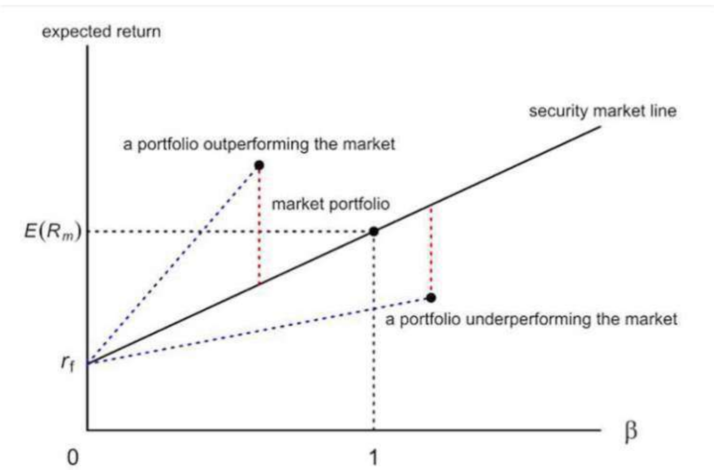

Figure 5.3: Example of the security market line [21]

This line demonstrates how all portfolios and assets behave in a market equilibrium compared to the market portfolio, i.e. that all investments do in fact reward the investor for their risk if they lie above this line or vice versa do not otherwise.

Let us see the return in formula 5.23 as the required return, that means how the market sees the expected return on a given security. Hereby, the market view would be an average of all market participants. One can imagine, that every single one of them has a sligthly different view and opinion on each asset. If one investor believes, that for a given security the true expected return is higher than the required return, i.e.

$$\mu_i > R + \beta_i(\mu_m - R), \tag{5.32}$$

then he believes that this asset is underpriced. That means that it would lie above the capital market line in figure 5.3. Conversely, if he believes that the security underperforms the expected return, then it would lie beneath it and would therefore be overpriced. In both cases, something similarly would happen. For underpriced assets, more and more investors would buy this specific asset, which would increase the price of it due to the risen demand. Overpriced assets would similarily get sold or even shorted, which would lead to a price drop. Due to this chain reactions, one should observe price adjustments to reach the equilibrium of the CAPM formula again. [18][p.67 ff.]

5.2.8 Conclusion

As a result, the capital asset pricing model provides a theoretical framework to estimate expected returns based on their risk factors and use those to calculate theoretically optimized portfolios. With this informations, investors can achieve a better understanding of the systematic and specific risk, that they take by investing in their portfolio. This allows them to make better informed decisions on purchasing and selling certain assets.

However, they also have to keep in mind the downsides of this model. As with all models, it cannot fit the whole truth into it and made some assumptions that may not hold in the real world. Additionally, it is only one of many tools on the way to prepare oneself for the financial world and should not be used as a sole information base.

5.3 The Entropy Model

One of the main problems, when it comes to mean-variance portfolio optimization, is that often the resulting portfolios are highly concentrated in a few assets. This stands in contrast to the diversification principle that lies beneath the whole idea of optimal investing. In general, this disadvantage can be traced back to the statistical errors, that occur when one estimates the inputs of the mean-variance model. As already mentioned, those errors then can lead to very extreme positions in the portfolio.

To prevent this to happen, one can use a different approach by measuring the diversity or randomness within a portfolio. This can be done by using the Shannon entropy, introduced in 1948 by Claude Shannon. Although, Shannon was an expert in information theory and his Entropy measure is of great use in physics as well, it found its way into the modern portfolio theory.

Another advantage of using entropy is that it always exists as opposed to the variance, which can in theory also expand to infinity.

This section aims to explain the theoretical foundations, methodology and practical implications of the incorporation of the entropy measure into the portfolio optimization process.

5.3.1 The Shannon Entropy

As mentioned before, Shannon had great expertise in the field of information theory and introduced the concept of entropy in conjunction with communication theory. The goal of it is to measure how much randomness is included in the distribution at hand. In other words, it represents the degree of uncertainty or disorder in an event. Another way to imagine it, is to see it as the average amount of information needed to describe the possible outcomes of a random variable.

Definition 5.3.1. Given a discrete random vector X with probability distributions $P = (p_1, p_2, ..., p_n)$ that satisfy the normalization condition $\sum_{i=1}^{n} p_i = 1.$, we can define the entropy in formula notation as

$$H(P) = -\sum_{i=1}^{n} p_i \log(p_i), \tag{5.33}$$

where p_i is the probability of event i and $\log(p_i)$ the information that it contains.

Theorem 5.3.2. *This entropy concept has three very important mathematical properties, which are very useful when using the mean-variance model.*

(a) The first such characteristic is the non-negativity.

(b) Another property of the Shannon entropy is that there exists a maximum and a minimum value. The maximum value is reached when $p_i = \frac{1}{n}$ for $i = 1, .., n$

$$H(P) = -\sum_{i=1}^{n} \frac{1}{n} \log(\frac{1}{n}) = -\log(\frac{1}{n}) \tag{5.34}$$

and the minimum value is achieved when exactly one $p_i = 1$ and all others are 0. Then the formula delivers $H(X) = -\log(1) = 0$. These optimums are also unique due to the next property.

(c) The last of the three characteristics is the strict concavity of the entropy function.

Proof. To prove *(a)*, one takes $0 \leq p_i \leq 1, \log(p_i) \leq 0$ and $-\sum_{i=1}^{n} p_i \log(p_i) \geq 0$. Then it can easily be shown with the definition of the natural logarithm that equality only happens when exactly one of the $p_i = 1$ and all others are equal to 0. In the context of the portfolio optimization process, the entropy is incorporated in the objective function to determine the portfolio weights. Therefore, the obtained weights are automatically non-negative, which implies that the entropy model assumes no short-selling at all.

The proof of *(b)* and *(c)* can be done by applying the definition of strictly concave on two different random vectors P and Q with the parameter $\omega \in (0,1)$. This delivers the inequality

$$\omega H(P) + (1-\omega)H(Q) \leq H(\omega P + (1-\omega)Q) \tag{5.35}$$

$$-\omega \sum_{i=1}^{n} p_i \log(p_i) - (1-\omega) \sum_{i=1}^{n} q_i \log(q_i) \leq -\sum_{i=1}^{n} (\omega p_i + (1-\omega)q_i) \log(\omega p_i + (1-\omega)q_i) \tag{5.36}$$

This can be transformed into

$$\omega \sum_{i=1}^{n} p_i \log\Big(\frac{\omega p_i + (1-\omega)q_i}{p_i}\Big) + (1-\omega) \sum_{i=1}^{n} q_i \log\Big(\frac{\omega p_i + (1-\omega)q_i}{q_i}\Big) \leq 0. \tag{5.37}$$

In this last inequality, the equality is only reached when $P = \omega P + (1-\omega)Q$, which would mean that P is exactly identical to Q. However, this contradicts the assumption that $P \neq Q$ and therefore the equality cannot be reached. Furthermore, we know for sure that the natural logarithm is concave and therefore can use the Jensen Inequality. This delivers the following

$$\begin{aligned} \sum_{i=1}^{n} q_i \log\Big(\frac{\omega p_i + (1-\omega)q_i}{q_i}\Big) &\leq \log\Big(\sum_{i=1}^{n} q_i \frac{\omega p_i + (1-\omega)q_i}{q_i}\Big) \\ &= \log\Big(\sum_{i=1}^{n} p_i\Big) \\ &= \log(1) = 0. \end{aligned}$$

Now we can use this result in equation 5.37 for both terms in the same way and see that the inequality stands since both ω and $(1-\omega)$ are greater than 0. This also concludes the proof of the strict concavity of the entropy function. [2] □

5.3.2 Relative Entropy and Cross-Entropy

There are also some other entropy measures beside the Shannon entropy, that can also be shown useful for portfolio optimization.

Theorem 5.3.3. *(Gibbs' inequality) To start off, consider two different, discrete probability distributions P and Q. The Gibbs inequality holds for all such distributions.*

$$-\sum_{i=1}^{n} p_i \log(q_i) \geq -\sum_{i=1}^{n} p_i \log(p_i) \tag{5.38}$$

In the case of continuous probability distributions, it is defined by using the probability density functions $p(x)$ and $q(x)$.

$$-\int_{-\infty}^{\infty} p(x) \log(q(x))dx \geq -\int_{-\infty}^{\infty} p(x) \log(p(x))dx \tag{5.39}$$

Proof. Let's start off with the discrete case. The proof again relies on the strict concavity of the natural logarithm and the Jensen inequality. For two probability distributions the following inequality holds.

$$\log\left(\sum_{i=1}^{n} p_i \frac{q_i}{p_i}\right) \geq \sum_{i=1}^{n} p_i \log\left(\frac{q_i}{p_i}\right) \tag{5.40}$$

On the left side, one can simplify the term and use that the sum over all q_i is equal to 1.

$$\log(1) \geq \sum_{i=1}^{n} p_i \log\left(\frac{q_i}{p_i}\right) \tag{5.41}$$

$$0 \geq \sum_{i=1}^{n} p_i \log(q_i) - \sum_{i=1}^{n} p_i \log(p_i) \tag{5.42}$$

The proof for the continuous case can be done analogously. □

Then we can define the cross-entropy between discrete distributions P and Q as

$$H(P,Q) = -\sum_{i=1}^{n} p_i \log(q_i) \tag{5.43}$$

and for continuous ones as

$$H(P,Q) = -\int_{-\infty}^{\infty} p(x) \log(q(x))dx. \tag{5.44}$$

One can show easily that equality only holds if P and Q are the same distribution. In the context of portfolio optimization, it measures the dissimilarity between the predicted returns of a portfolio and the observed ones and it is therefore often used as a loss function.

To go further into detail, we can look at the difference between the cross-entropy and the already introduced Shannon entropy.

$$KL(P,Q) = H(P,Q) - H(P) \tag{5.45}$$

$$= \sum_{i=1}^{n} - \sum_{i=1}^{n} p_i \log(q_i) + \sum_{i=1}^{n} p_i \log(p_i) \tag{5.46}$$

$$= \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right) \tag{5.47}$$

This is called the relative entropy or also the Kullback-Leibler divergence. It compares the chaos or uncertainity in the two distributions or, in other words, it is a measure of how one probability distribution diverges from a second one. The continuous version again looks like this:

$$KL(P,Q) = \int_{-\infty}^{\infty} p(x) \log\left(\frac{p(x)}{q(x)}\right)dx \tag{5.48}$$

However, if the distribution P is known, the relative and cross-entropy can be viewed equivalently since the entropy of P becomes a constant and the only difference between the two measures. Therefore, we can choose one of them for our optimization model going forward. [22]

5.3.3 Entropy Model

The model itself is quite similar to the basic Markowitz model itself. With the same assumptions taken in chapter 5.1, one can adapt the markowitz model (5.7-5.9) to include the Shannon entropy of the portfolio P with n assets itself.

$$min\ \mathbb{V}(R) - \lambda H(P)) \tag{5.49}$$

$$s.t.\quad \mathbb{E}[R] \geq \mu \tag{5.50}$$

$$and\quad \sum_{i}^{n} p_i = 1\quad with\ 0 \leq p_i < 1,\ \ i = 1, 2, ..., n. \tag{5.51}$$

Again, p_i stands for the weight of the i-th asset in the portfolio and R is the aggregated return. Additionally, there is the real constant λ, which is chosen according to the risk appetite and strategy that the investor wants to pursue. It affects the impact of the entropy on the optimization. The idea of this concept is to minimize the variance and to maximize both the return and the entropy at the same time.

Theorem 5.3.4. *The solution to the minimization problem 5.49 is unique.*

Proof. Firstly, it is necessary to show that the following function is convex.

$$f(x) = \lambda(\mathbb{V}(R) - \zeta H(P)) - \mathbb{E}[R] \tag{5.52}$$

Since the covariance matrix is positive definite, the first term is convex as a function of the portfolio weights for $\lambda > 0$. The entropy is a concave function as shown above. This leads to the second term being convex, if both $\lambda > 0$ and $\zeta > 0$. The last term is linear as a function of the weights and therefore does not impact the overall convexity. All in all, the function is convex and therefore any local minimum is a global minimum. Furthermore, there only exists one such minimum within the feasible region, since this region is also convex due to the constraints. [2] □

5.3.4 Entropy Transformation

As already described in the introduction, companies may have different goals to reach. They might have subportfolios in their portfolio where they try to achieve different levels of return with different levels of risk as well. For instance, insurance companies have so called fund-linked life insurance products. Within those subportfolios, the percentage of equities can vary, since they are both the riskiest as well as the most profitable assets in general.

So how can this be incorporated in the optimization model? We propose to incorporate the Kullback-Leibler divergence, already described in (5.45-5.47). Replacing the entropy in (5.49-5.51) with the KL divergence should do exactly what we want.

$$min\ \mathbb{V}(R) - \lambda KL(P, Q)) \tag{5.53}$$

$$s.t.\quad \mathbb{E}[R] \geq \mu \tag{5.54}$$

$$and\quad \sum_{i}^{n} p_i = 1\quad with\ 0 \leq p_i < 1,\ \ i = 1, 2, ..., n. \tag{5.55}$$

This way, an investor can tell the model, what preferences he has regarding the asset weights and the model tries to optimize the weights with respect to that. This increases the flexibility of the optimization process immensely since the investor is not constrained to a uniform distribution of his weights. It enables the optimization of benchmarks in regard to risk tolerance by incorporating investment strategy or market views. Furthermore, this model still tries to make a more stable and predictable portfolio performance as extreme allocations are discouraged and the weights seen over time tend to be smoother than with convential mean-variance analyses. The uniqueness of the solution follows with theorem 5.3.3 and the convexity of the KL divergence.

5.3.5 Adapted Entropy Model

The choice for the parameter λ in the previous model could turn out to be quite problematic since the market can be very dynamic and require it to change over time. To achieve such an adaption, one could include λ as a variable depending on different observations rather than a pure constant.

$$min \quad \mathbb{V}(R) - \lambda_t H(P) \tag{5.56}$$

$$s.t. \quad \mathbb{E}[R] \geq \mu \tag{5.57}$$

$$and \quad \sum_i^n p_i = 1 \quad with \ 0 \leq p_i < 1, \ i = 1, 2, ..., n. \tag{5.58}$$

This definition is equivalent to the one in the standard entropy model (5.49-5.51) with the addition of λ introduced as a stochastic process. The big question is how to model said dynamic parameter λ based on the historic data. In this thesis, we try out different approaches to see what would work best.

Empirical Testing

This method is probably the most straight-forward one there is. We plug in values from 0 to 1 with steps of 0.01 for our λ into the entropy model (5.49-5.51) and optimize the portfolio. This way the effect of the parameter is directly observable. Then based on what λ produces the highest realized Sharpe ratio in the sample, we decide to take that. Although, this method takes up the most resources since there are more optimization runs to be done in each time step.

Performance Based λ

Let's take time steps to be one month and the vector of the portfolio weights belonging to the previous month $t-1$ is $P_{t-1} = (p_{(1,t-1)}, p_{(2,t-1)}, ...p_{(n,t-1)})$. Analogously, the return of the time step is in the vector R_{t-1} and the expected return of each asset i is also put into the according vector $M_{t-1} = \mathbb{E}[R_{t-1}]$. Additionally, we also have the $n \times n$ covariance matrix of the past step $\mathbb{V}_{t-1} = \mathbb{E}[R_{t-1} - \mathbb{E}[R_{t-1}]]^2$.

Suppose a starting value λ_0. With this starting value, we can optimize the first portfolio. From there,we can calculate a value for each time step t by using the following formulae

$$\mu_t = P_{t-1}^T R_t.$$

This corresponds to the actual return in the past month. Now, we can compare this figure to our target return μ. Depending on that, we can adjust our λ in the following ways.

- **Performance risk-affine**: If we failed to meet our target return, i.e. $\mu_t < \mu$, we try to increase the parameter.

  $$\lambda_t = \lambda_{t-1} + \mu_t \tag{5.59}$$

  The idea of this is, that the optimization penalized the risk of the assets too much. Since risky assets come with more volatility but also more possible gains, the risk term in the optimization tends to give them smaller weights than. Therefore, increasing the parameter leads too less focus on the risk and we might increase the performance for the next term.

  If the return exceeds our target, the investor might want to decrease the risk further and secure his gains. Due to this, we can decrease the λ:

  $$\lambda_t = \lambda_{t-1} - \mu_t$$

- **Performance risk-averse**: What if the investor thinks exactly the other way around? His idea might be that when the target is not met, the risk was too high. Consequently, he wants to decrease the λ to get more emphasis on the risk term.

  $$\lambda_t = \lambda_{t-1} - \mu_t \tag{5.60}$$

  Of course, if the target is met, the investor might seek to increase the λ since the markets are going good in his opinion and the need for risk-aversion is not so high.

  $$\lambda_t = \lambda_{t-1} + \mu_t$$

As a result, this model can adapt to the different effects that can be observed on the financial market by adding this dynamic parameter λ_t due to the fact that it considers also the diversification of the portfolio and thus creates a more robust one. The question is now which way the optimization works the best. [2]

6 Test Data

The data used to test these models was provided by the Niederösterreichische Versicherung AG. Specifically, the optimization was done with four different asset classes, each represented by an according index. The data was provided on a monthly base from 31.12.1998 on until 30.09.2024. The four indices are shortly described in this section.

6.1 Global Equities

The MSCI World index is one of the most famous indices out there. It tracks the performance of companies from large to mid-size market capitalisation in the developed world. Approximately 85% of the free float-adjusted market cap is covered in each of the countries, that are monitored by this index. There is no restriction to any sectors, which provides a broad representation of the global equity market.

6.2 Government Bonds

The next index is one to measure fixed-income investments. More specifically, it is designed to track the performance of Euro-denominated government bonds with maturities over one year. Governments in the Eurozone tend to have a higher level of credit quality, which represents their low risk of default and high stability. However, this also leads in general to lower coupons than for other fixed-income investments, which was observable in the last decade, where the interest rates were close to zero. All in all, it provides a comprehensive and reliable benchmark for government debt performance inside the Eurozone.

6.3 Corporate Bonds

Another component of the money market are corporate bonds. The index used to track them consists of fixed-rated bonds issued by a diverse range of corporate entities, including the industrial, utility or financial sector. All of them are denominated in Euro to avoid currency risk within the index. To be included, they have to have an investment-grade, which ensures a higher credit quality and leads to lower default risk. To sum up, this index delivers a transparent and objective benchmark for the performance of liquid investment-grade corporate bonds in the Eurozone market.

6.4 Emerging Market Bonds

Since the previous indices have more of a focus on Europe or other developed countries, the last one aims to track a different class of bonds. It tracks the performance of corporate bonds issued in hard currencies by companies in emerging markets such as China, Brazil or South Africa. The requirement of an investment-grade is similar to the previous ones, making it also less risky than benchmarks with not rated bonds. To mitigate currency risk, the index also includes a hedging mechanism, ensuring more stable returns while also strengthening the comparability with the other indices used. Again, a broad range of different sectors are tracked, delivering a good overview of corporate bonds in emerging markets.

7 Results

The models were all implemented using the software R. First off, the three described model - Mean-Variance (MV), Entropy (E) and Adapted Entropy (AE) were tested without a target return. This means, they solely produce their respective minimum variance portfolio. It is important to note, that here the λ in the Adapted Entropy model was always chosen by the empirical testing method, since the model lacked a target to beat in terms of performance and therefore the decision on how to proceed in the other methods. For the standard entropy model, the test included λ in the range of $[0, 1]$ to see some differences there as well. From thereon, the target return was introduced to create the efficient frontier in respect to each of their risk measures.

After this first stage of testing, three fictive benchmarks were introduced to see which model performed the best in regard to those. For the three standard models, weight boundaries were added to make them comparable to the benchmark. Additionally, the entropy transformation was tested for both the Entropy and the Adapted Entropy model to see, how the weights and performance evolved with this change. Moreover, the different methods on how to choose the λ in the Adapted Entropy were tested alongside, to also monitor their impact on the optimization.

7.1 Standard Optimization

To start things off, we did the optimization without a return target to get a feeling for how the various models behave. As already mentioned, the data is available from the year end of 1998 so we start there. To estimate the covariance as well as the expected returns, we take the first 10 years as our sample, therefore the start of the optimization is the 31.12.2008. Then, the portfolios are calculated on a monthly basis. This way, the weights can be adjusted on a monthly basis based on new information. The sample data for the parameter estimation is also moving along over time. In other words, for every time step, that is added to the sample, the oldest data is deleted to have a bigger focus on more recent data.

7.1.1 Performance

When simply minimizing the respective risk, the entropy model generated a higher performance compared to both the mean-variance and the adapted entropy model due to its approach in weighting equities as can be seen in table 7.1.1 and figure 7.1. The key reason here is that the entropy model does not avoid the risk of equities as much as the other two models, since it is more focused on the uniform distribution between the asset classes. By doing so, it captures a broader range of market movements. In other words, this model seems to perform better as the impact of any single asset's poor performance is mitigated

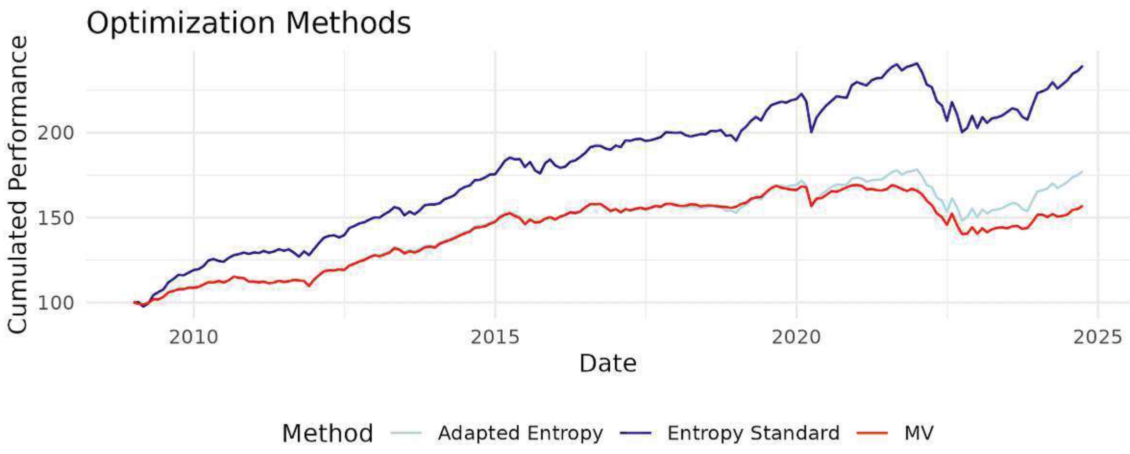

Figure 7.1: The cumulated performance of the three optimization methods from 2008 until today

by the equal presence of other assets in the portfolio. The volatility however is also higher compared to the other two models. This is not surprising as the main focus in this model is not on the risk. The Sharpe ratio on the other hand is also increased. The entropy portfolio therefore seems to return a better trade-off between its risk and return making it more appealing.

Model	Return p.a.	Vola p.a.	Entropy	Sharpe Ratio
Mean-Variance	2,89%	4,58%	0,46	63,04%
Entropy	5,67%	5,90%	1,39	70,33%
Adapted Entropy	3,68%	5,24%	1,39	96,02%

Table 7.1: Key figures for the three models in the standard optimization run

In contrast, the mean-variance model focuses on minimizing the risk. While effective, this model can - as already mentioned - sometimes lead to portfolios that are overly concentrated in assets with historically high returns and lower variance, potentially increasing risk if market conditions change.

The adapted entropy model, on the other hand, underperformed primarily due to its choice of lambda, a parameter that controls the trade-off between entropy and volatility. For most of the time, the lambda calculated by empirical testing was close to zero, which means the model leaned heavily towards the mean-variance. This resulted in portfolios that participated less in the increases in the market.

The evolution of the parameter λ can be seen in figure 7.2. Two time steps stand out to the others. At the start the mean-variance model always returns a portfolio with a higher Sharpe ratio than the adapted entropy model until September of 2018. In this specific time step, the algorithm cuts off a month, where the emerging market bonds performed exceptionally bad with a loss of over 20%. As already observed, the mean-variance optimization does not include this asset class at all. This leads to a jump in the return and volatility of this asset which is also used for the calculation of the Sharpe ratio and consequently the λ. Therefore, the entropy slightly starts to get involved up until July of 2021. Here, the same argument can be stated for the emerging market bonds as well as equities. However, in this time step the situation gets turned upside down. The λ jumps to 1 and stays there until the rest of the timeline. This brings up the question, if this method is too unstable and sensitive to use. This is gonna be discussed in section 7.3 together with the other methods to calculate it.

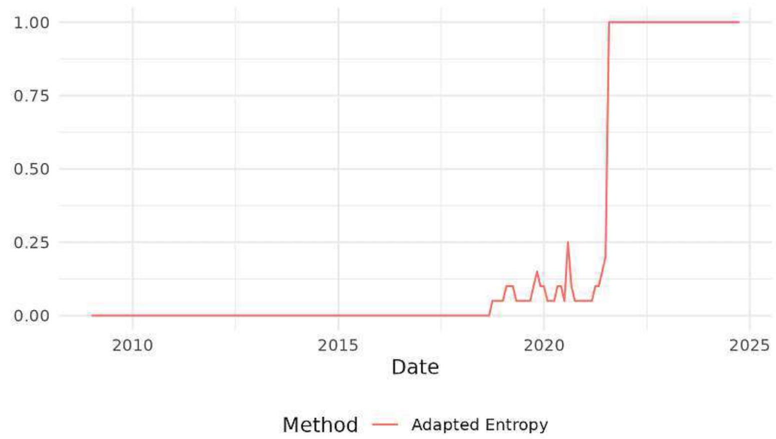

Figure 7.2: Choice of λ by empirical testing over time

7.1.2 Weight Evolution

The evolution of the weights for each portfolio can be seen in figure 7.3. One has to notice the breaks of the y-axis for each model. The mean-variance and adapted entropy model both have quite a big range, while the entropy model never really deviates much from the 25% mark, since there is maximum entropy reached. Here, one can see perfectly that the AE portfolio tracks the MV one almost exactly up until the year 2019. What happens then is that the MV model invests even more in corporate bonds, which makes it even less appealing. Therefore, the λ parameter changes very quickly from 0 to 1 and it tracks the entropy version.

Another advantage of the entropy optimization can also be observed, when looking at these graphs. While the mean-variance has big changes in its portfolio and even excludes one asset class completely, the entropy model moves rather smoothly. The size of the jumps and the evolution itself is way more robust than in the former. In reality, this would be also the goal of an investor since very few of them want to reallocate their whole portfolio on a monthly basis, but make rather small changes to it to generate more stability.

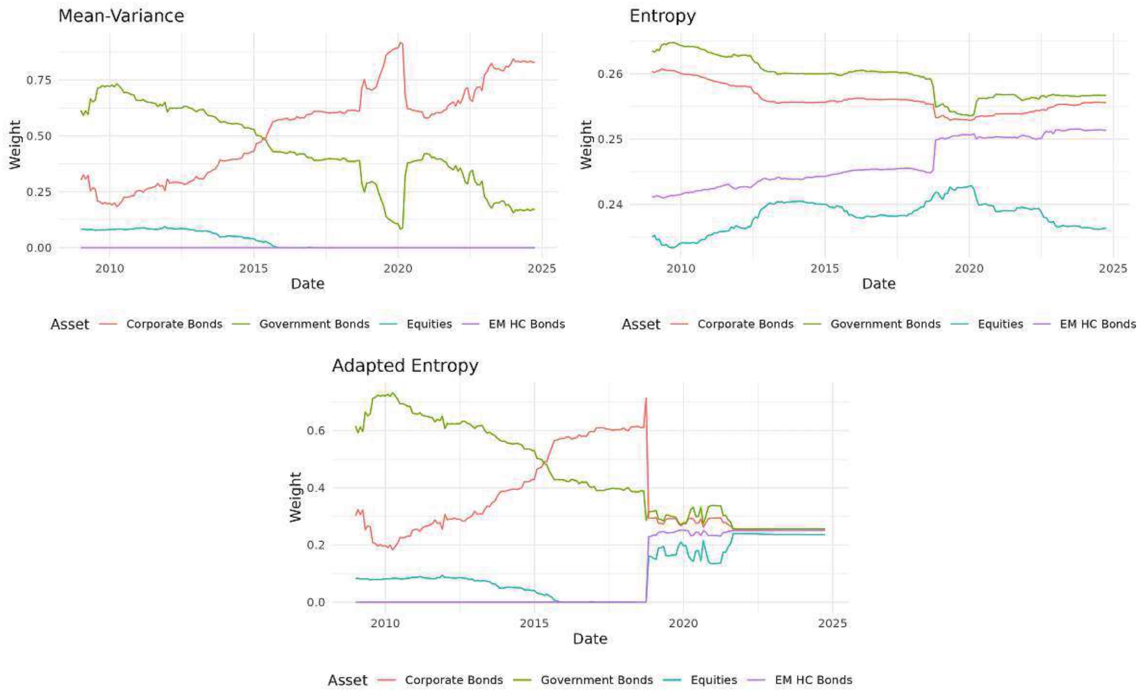

Figure 7.3: Weights of the asset classes for each optimization model over time

7.1.3 Efficient Frontiers

Moving on, we also visualized the efficient frontier in regard to the two different risk measures, the volatility and volatility reduced by the entropy. This is necessary due to the efficient frontier changing according to which objective was pursued in the optimization. The graphs on display respond to the last timestep in the optimization process since it is always drawn for one time step. Consequently, the adapted entropy portfolio was left out since in the last time step $\lambda = 1$ and therefore its efficient frontier was the same as for the entropy model.

What can be observed in the graph 7.4 is, that in regard to the volatility, the mean-variance optimization always performs better. The entropy model delivers a bit of a bewildering graph. The maximum return portfolio for both models is the same, since this corresponds to the asset class with the highest expected return. Both algorithms need to invest 100% in that to reach the most return. This portfolio however is not the one with the highest volatility in the set of optimal ones returned by the entropy optimization.

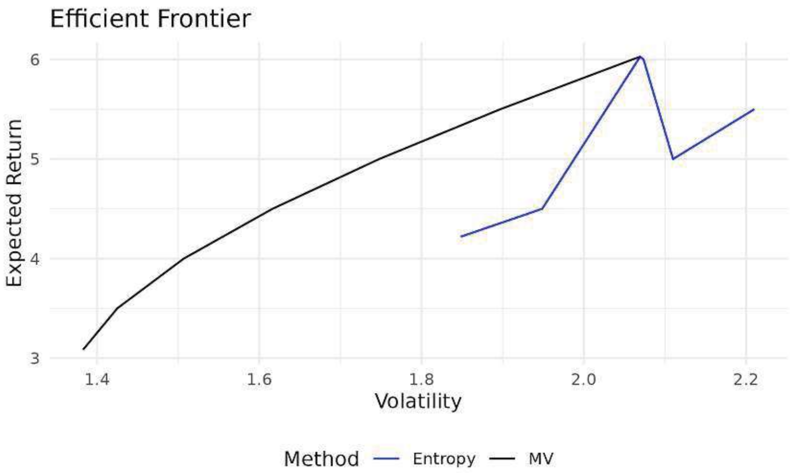

Figure 7.4: The efficient frontier in regard to volatility

However, as soon as we change the risk measure and subtract the entropy from the volatility, we end up with figure 7.5. There, one can see that the mean-variance model delivers a rather wild graph up until an expected return of 5%. From there the deviance is shrinking due to the fact that the impact of the entropy is shrinking since both models need to invest in asset classes with higher expected returns. Figure 7.6 displays the evolution of

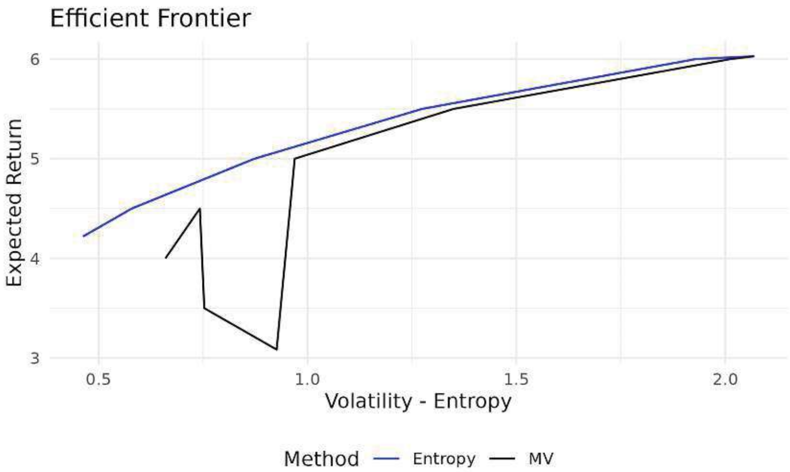

Figure 7.5: The efficient frontier in regard to volatility reduced by the entropy

the asset weights for the portfolios on the efficient frontier. It emphasizes the statement above. With a low target return, the entropy model on the right can equally distribute the portfolio, but the higher the goal is, the more it has to deviate from that as only emerging markets have an expected return over 6% at this point in time.

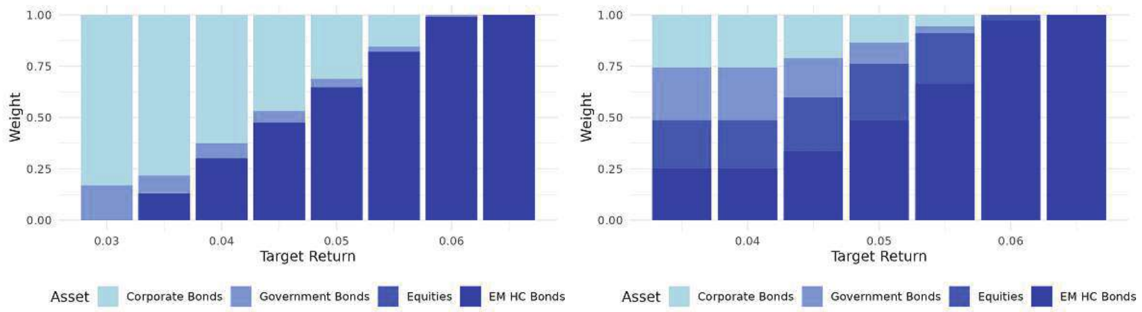

Figure 7.6: Weights of the asset classes of the optimal portfolios - on the left for the mean-variance model and on the right for the Entropy model

To put it in a nutshell, in the first tests we saw that the entropy model performs quite well compared to the mean-variance model, although the volatility increases when investing in an uniform way. Nonetheless, the Sharpe ratio was higher in the entropy model making it the best in terms of risk-return trade-off. Therefore, it seems quite appealing for an investor to distribute his capital equally.

7.2 Optimization of Benchmarks

The next step in the optimization is to see how the models can perform, when trying to track a certain benchmark. For this purpose, three fictive sets of index weights were constructed - conservative, balanced and dynamic. These names relate to the weight of the equity and emerging markets in the set. The more these asset classes are involved, the riskier - or dynamic - the portfolio becomes. The exact weights are displayed in table 7.2 below.

Benchmark	Government Bonds	Corporate Bonds	Equities	Emerging Market Bonds
Conservative	70%	10%	20%	0%
Balanced	40%	15%	40%	5%
Dynamic	10%	20%	60%	10%

Table 7.2: This table shows the weights of the asset classes for each benchmark.

Moreover, the different models were given an additional constraint to make them comparable. The target return μ was calculated as the expected performance of the benchmark in each time step. Since the weights of that index set do not change, this measure can be calculated easily. Moreover, this ensured, that all of the models had the same constriction to build something profitable.

7.2.1 Conservative Benchmark

To start off, we have a look at the supposedly least volatile and safest of the three weight sets. To the three already tested models were another two added. Both of the entropy and adapted entropy got transformed as described in section 5.3.4 so that their aim is not equally distributed weights, but to distribute the capital like the benchmark itself. In the following, these two are called entropy to benchmark and adapted entropy to benchmark. However, both still try to minimize the variance as well.

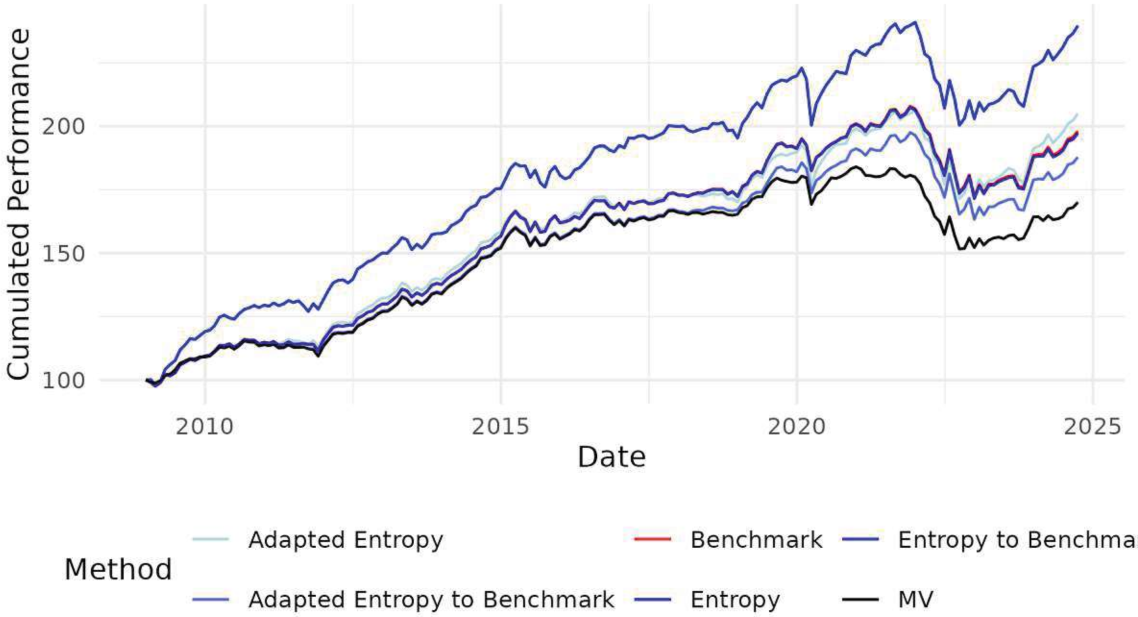

Figure 7.7: Performance of all five models compared to the conservative benchmark

As one can see in figure 7.7, the entropy portfolio outperforms the rest of the models by quite a margin in regard to the return. This makes sense since here both of the riskier asset classes emerging market bonds and equities are more included than in the other asset classes. Interestingly, this portfolio performed exactly as in the standard optimization run. That means that it actually can achieve the expected target return with its optimal portfolio in regard to its risk measure. The adapted entropy on the other hand also returns a better performance than the benchmark. Both of those methods also achieve a higher Sharpe ratio than this fixed index set as one can see in table 7.2.1 below.

The benchmark itself has only very limited focus on equities and excludes the former entirely so it is not surprising, that it misses that much performance. The entropy model that used the Kullback-Leibler divergence instead of the Shannon entropy tracked the benchmark almost perfectly. However, the adapted entropy model with the same transformation did have less performance. This is also understandable, since it tends to track the mean-variance, if the Sharpe ratio of that is better and ends up between the benchmark and the Markowitz model.

Model	Return p.a.	Vola p.a.	Entropy	Sharpe Ratio
Balanced Benchmark	4,41%	5,30%	0,80	83,21%
Mean-Variance	3,41%	4,66%	0,34	73,23%
Entropy	5,67%	5,90%	1,39	96,21%
Entropy to Benchmark	4,38%	5,28%	0,83	82,96%
Adapted Entropy	4,64%	5,31%	1,39	87,35%
Adapted Entropy to Benchmark	4,06%	5,08%	0,83	79,93%

Table 7.3: Key figures for the models compared to the conservative benchmark

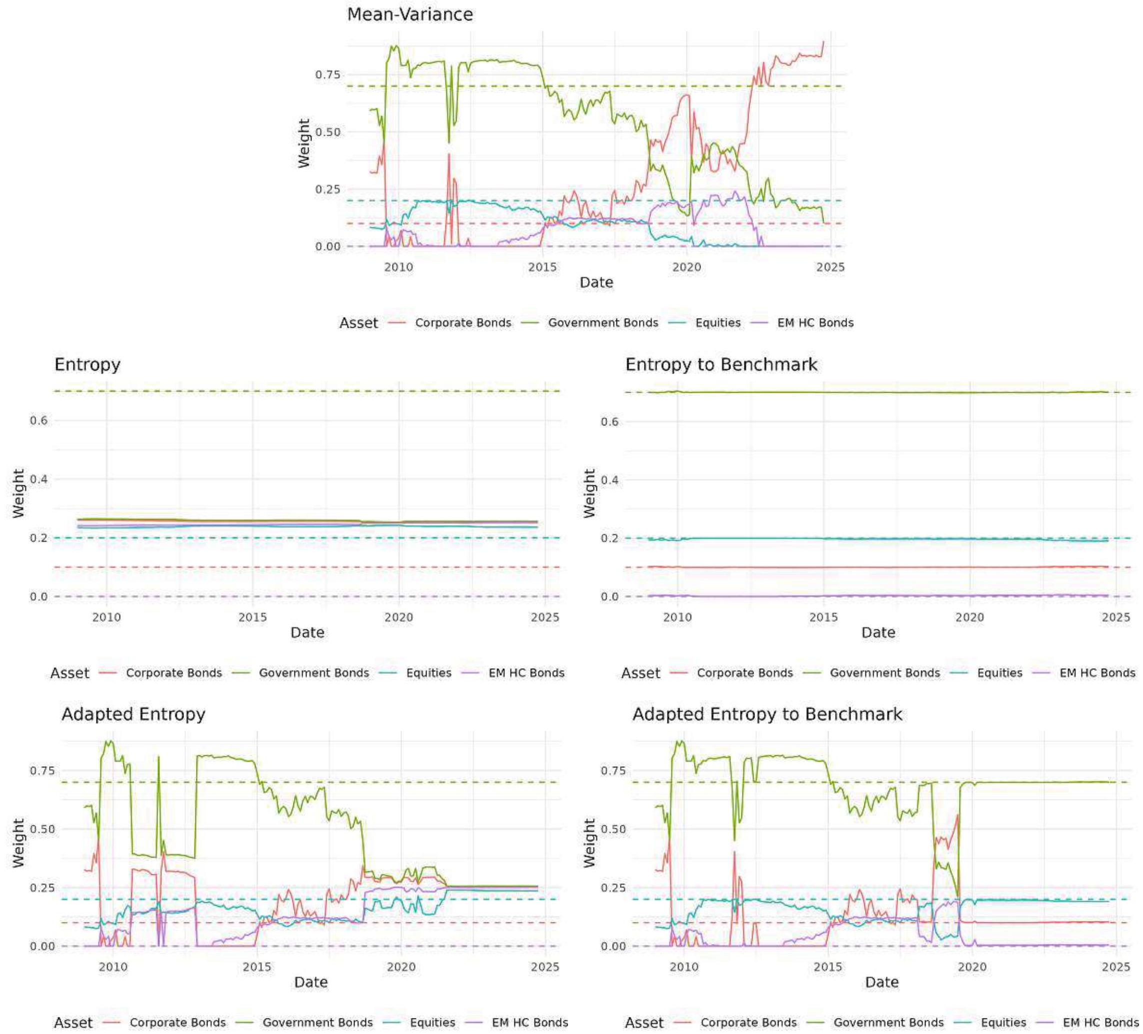

Figure 7.8: Weights of the asset classes of the optimal portfolios for the conservative benchmark

The weights of each of the optimized portfolios is visualized in figure 7.8. As already said in the previous section, the weights of the mean-variance portfolio tend to deviate a lot and also overly concentrate in specific asset classes whereas the two entropy portfolios almost perfectly track their respective equilibrium. The question is here, if the entropy gets too much attention or the target return is too low, but more in the analyses of the balanced and dynamic optimization.

The adapted entropy portfolios do tend to jump also quite. The calculation of λ is not really smooth so they change from the entropy to the mean-variance method very fast. In reality, no company would want to overhaul their portfolio in such fast and dramatic way.

7.2.2 Balanced Benchmark

Next up is the index set, that has a more balanced approach between risk and return. Again, all five of the portfolio models were calculated with the goal of reaching the same expected return in each time step as the balanced benchmark.

The picture already turned here, where no portfolio could outgain the benchmark. The adapted entropy with the KL divergence again stayed very close to it. All of the others failed to get as much as return as the two as one can see in table 7.2.2. The Sharpe ratios however are all quite close to each other except for the mean-variance one, which could indicate that those are still very risk efficient except for the latter.

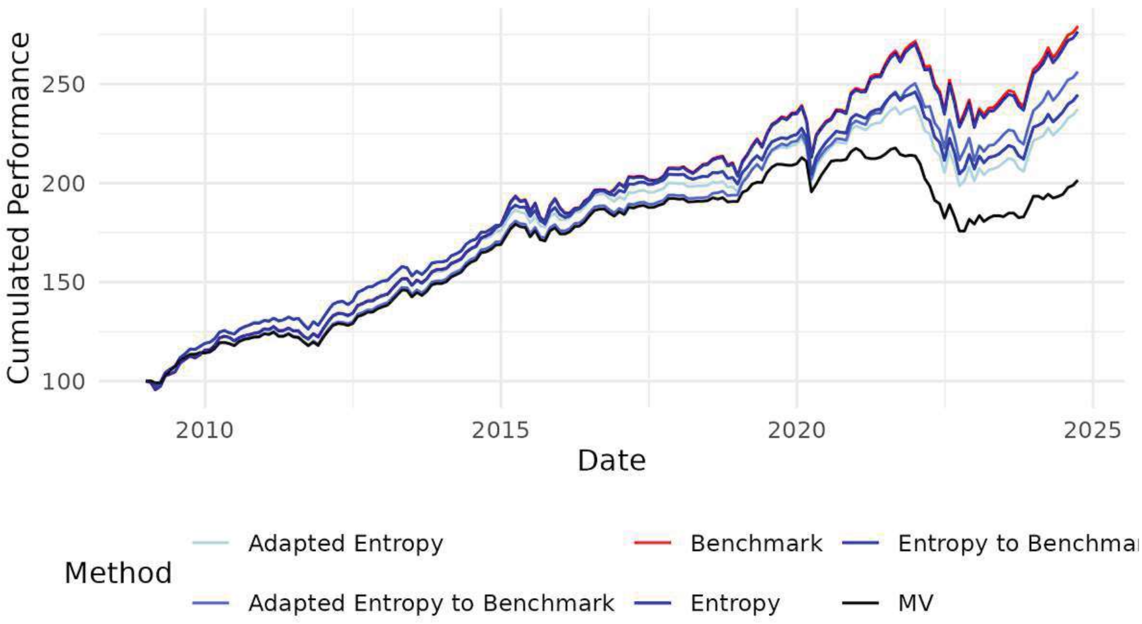

Figure 7.9: Performance of all five models compared to the balanced benchmark

Except for the transformed entropy model, all of them had to make adjustments to their weight distribution to reach the excepted return. The mean-variance had to give more focus to equities as well as emerging market bonds to get more return but excluded corporate bonds for a long time in respond to that. The entropy optimization had to deviate quite strongly compared to the tests before. In other words, if the expected return target rises, it has to deviate more from its equilibrium.

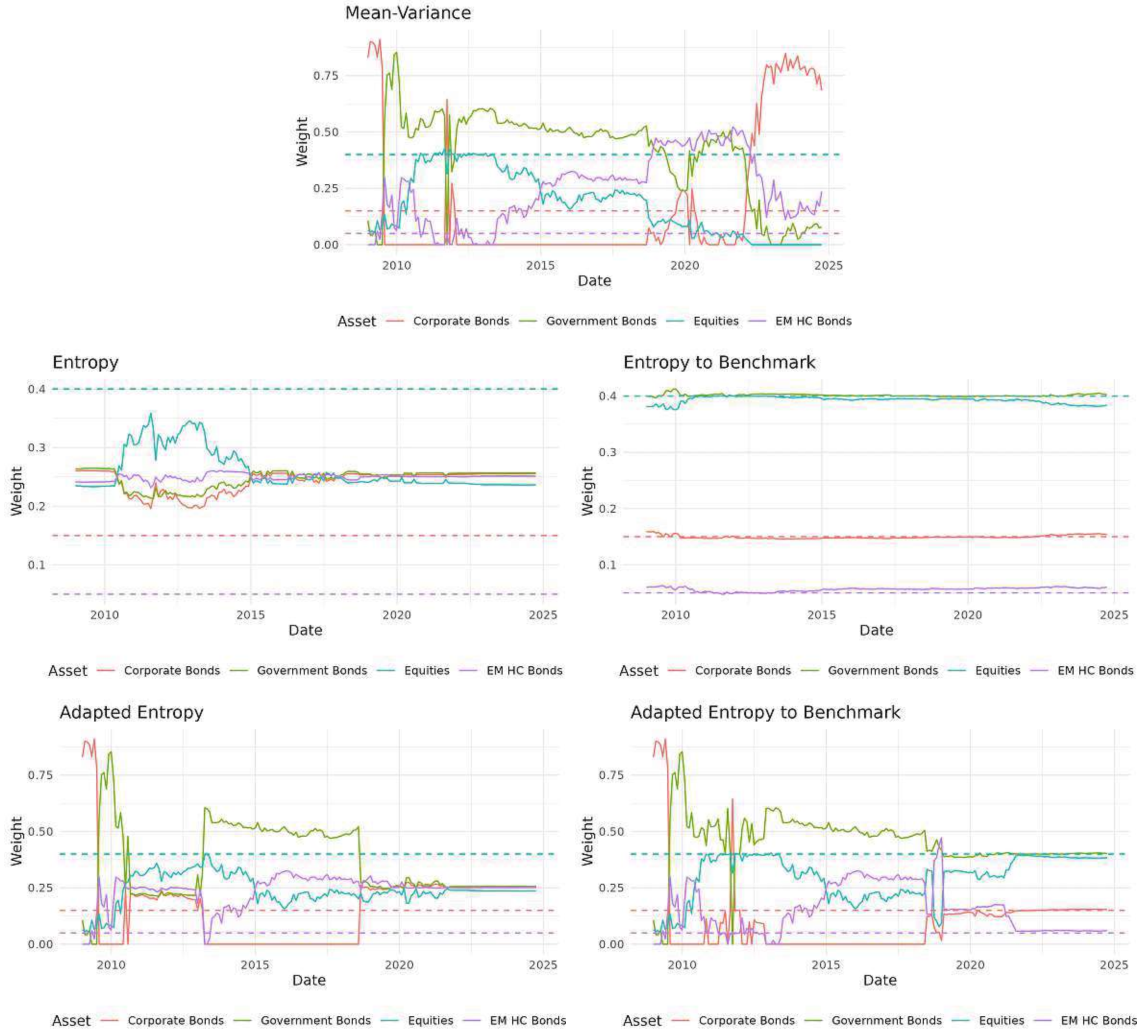

Figure 7.10: Weights of the asset classes of the optimal portfolios for the balanced benchmark

Model	Return p.a.	Vola p.a.	Entropy	Sharpe Ratio
Balanced Benchmark	6,70%	6,72%	1,16	99,66%
Mean-Variance	4,52%	5,22%	0,80	86,68%
Entropy	5,81%	5,97%	1,39	97,39%
Entropy to Benchmark	6,63%	6,66%	1,19	99,58%
Adapted Entropy	5,61%	5,73%	1,39	97,99%
Adapted Entropy to Benchmark	6,12%	6,02%	1,19	101,76%

Table 7.4: Key figures for the models compared to the balanced benchmark

7.2.3 Dynamic Benchmark

Last but not least, we have a look at the optimization with the highest percentage of equities included. The trend of the last section continues as the benchmark soars with the transformed entropy method a close second. What is really noticeable through these three test runs, is that the Sharpe ratio of the adapted entropy optimization with the KL divergence as objective surpassed all others as one can see in table 7.2.3. Meanwhile, the portfolio produced by the Markowitz model struggles to generate more return as well as a higher Sharpe ratio.

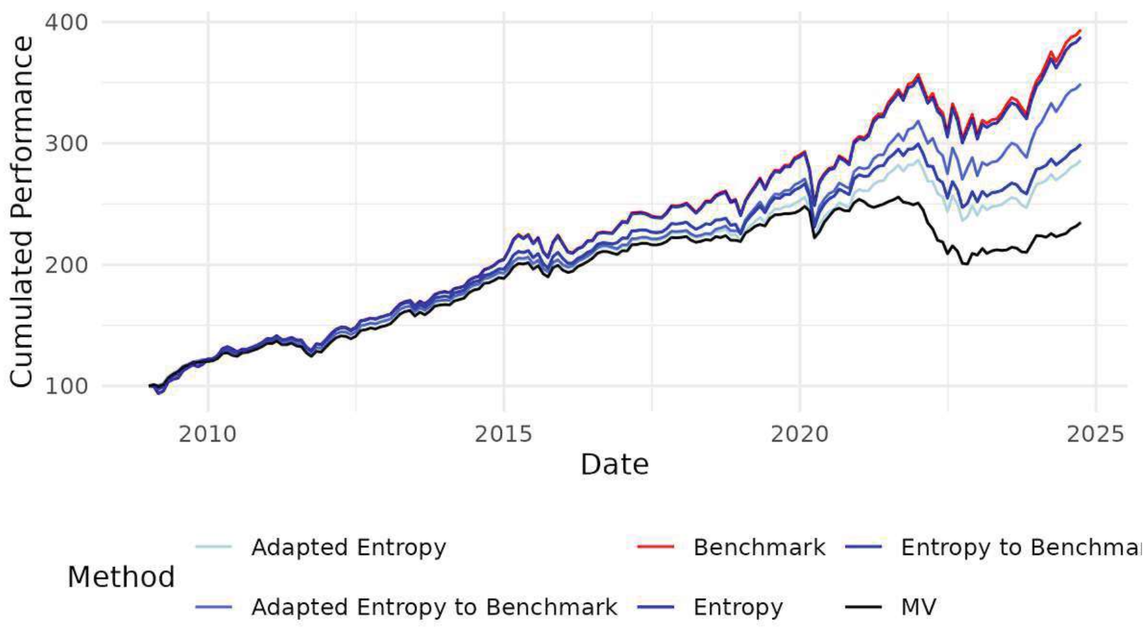

Figure 7.11: Performance of all five models compared to the dynamic benchmark

Model	Return p.a.	Vola p.a.	Entropy	Sharpe Ratio
Dynamic Benchmark	9,04%	8,81%	1,09	102,62%
Mean-Variance	5,54%	6,27%	0,87	88,29%
Entropy	7,17%	7,28%	1,36	98,50%
Entropy to Benchmark	8,93%	8,70%	1,13	102,67%
Adapted Entropy	6,86%	6,91%	1,36	99,25%
Adapted Entropy to Benchmark	8,21%	7,44%	1,13	110,36%

Table 7.5: Key figures for the models compared to the dynamic benchmark.

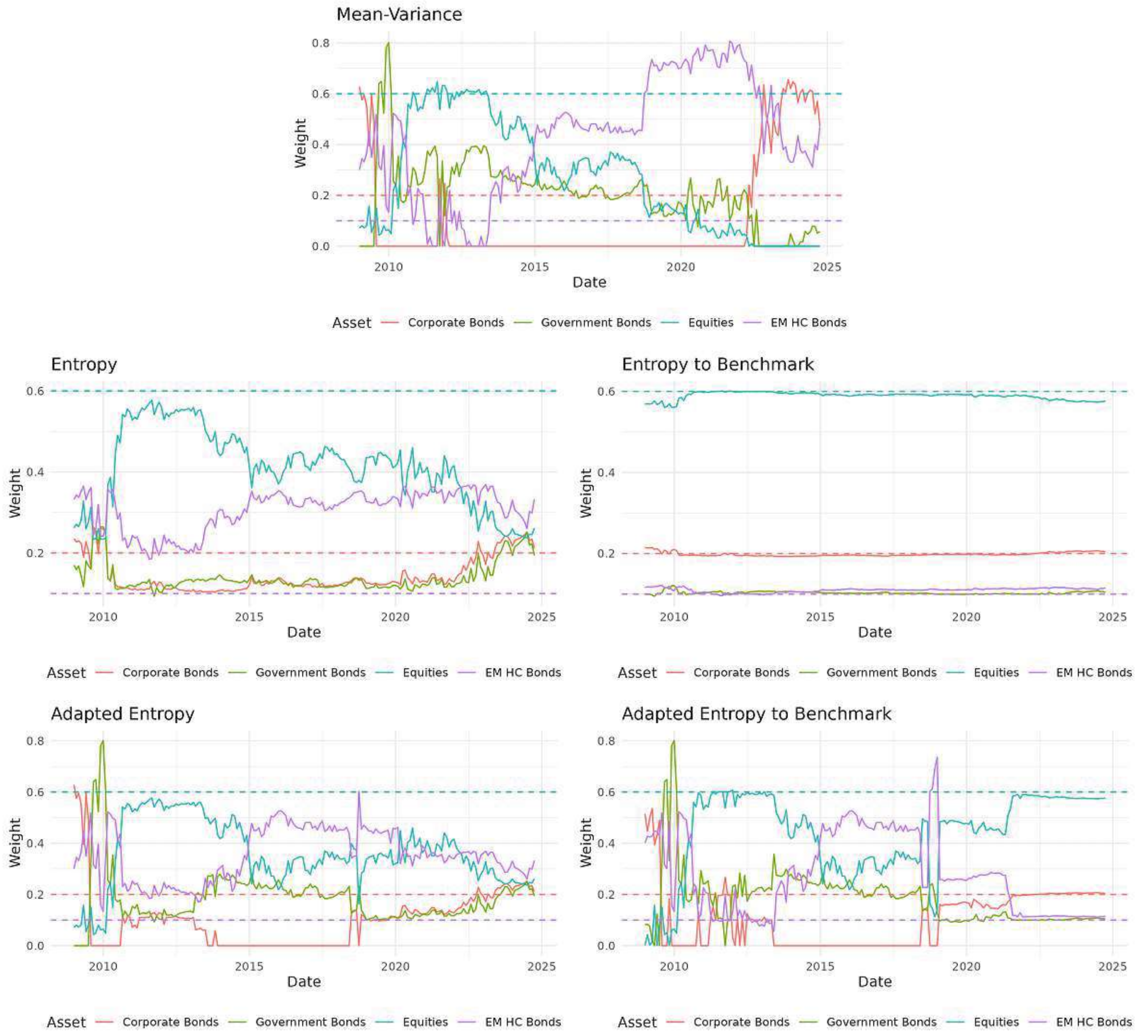

Figure 7.12: Weights of the asset classes of the optimal portfolios for the dynamic benchmark

The weights do continue the trend that we saw earlier. The entropy model has to readjust its weights on a regular basis and is far from an uniformely distributed portfolio. In general, all of the models take on a much higher investment in equities than before. The stakes in emerging markets also continued to rise, while both the government and corporate bonds lost their impact in the portfolios.

7.2.4 Conclusion

To sum up, we saw that the entropy portfolio does bode quite well. In every optimization, it returned a stable Sharpe ratio of close to 100%. In other words, this model achieved a very risk efficient portfolio, although the higher the target gets or the more the benchmark deviates from the uniform distribution, the more it has to vary its weights.

The adapted entropy method on the other hand did what we expected from it. It returned a portfolio that was between the benchmark and the mean-variance portfolios. This way, it always provided an overproportionate increase of return compared to the additional volatility to the mean-variance portfolio.

The change from the Shannon entropy to the Kullback-Leibler divergence seemed to perform especially well, when used in combination with a higher risk appetite and the adapted entropy model. In the last test, it scored the highest Sharpe ratio. Compared to the entropy model, it increased the return per anno over 1% while only taking on 0,16% more volatility.

7.3 Performance-Based Parameter Calculation

The last open question is the change to different methods to calculate λ. The model on its own performed rather well. The evolution of the parameter however is quite sensitive as can be seen in figure 7.13, where its change over the time is visualized.

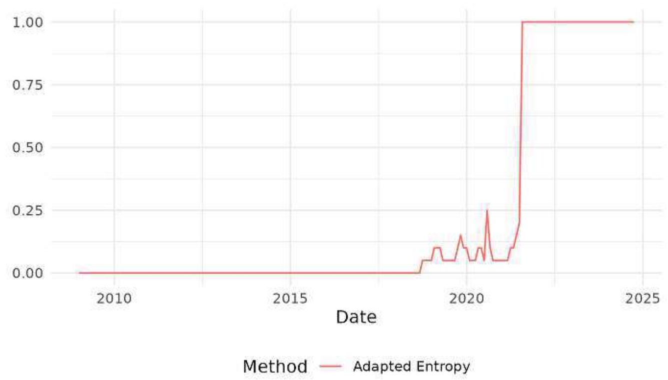

Figure 7.13: The evolution of λ in the adapted entropy model over the time

Therefore, we tried to adapt the way of the calculation as described in section 5.3.5. This was tested using the dynamic benchmark since there the adapted entropy portfolio with benchmark transformation scored the highest Sharpe ratio.

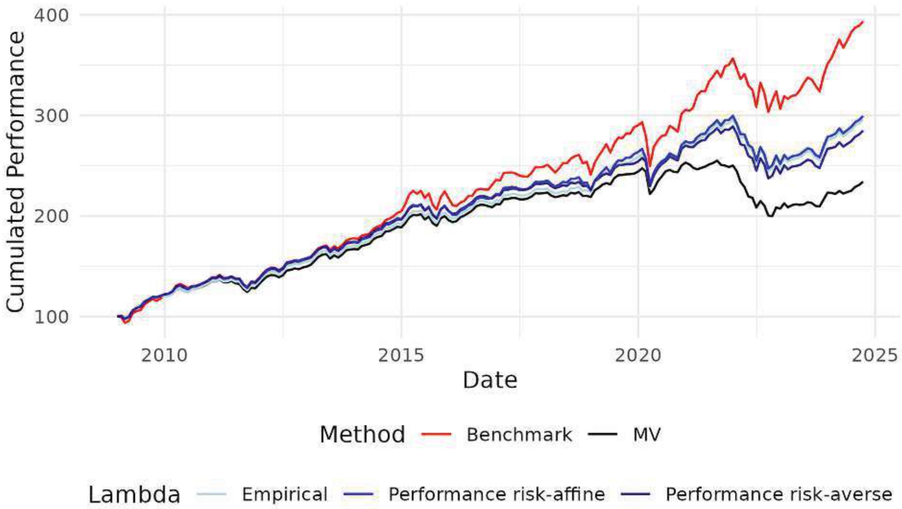

Figure 7.14: The performance of the adapted entropy model for different λ calculation methods

In the standard adapted entropy case, the different models however delivered no real change to the performance, which can be observed in figure 7.14. While the approach of increasing λ when the return target is not met, i.e. the performance risk-affine method, had a minimally higher return than the empirical version, its volatility rose more so that the Sharpe ratio went down. The other direction had a lower volatility as well as a lower performance, with a Sharpe ratio of just over 100%. Therefore, no real improvement had been achieved in this case although the evolution of λ did differ quite a lot, as can be seen in figure 7.15.

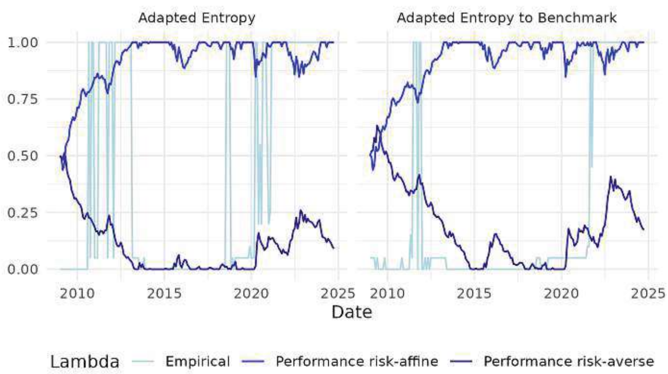

Figure 7.15: The evolution of λ in the different models

On the other hand, the spread of the results was higher in the version with the Kullback-Leibler divergence. Both versions of the performance-based λ calculation increased the performance of the portfolio. However, the volatility in both cases rose disproportionately leading to a shrinking Sharpe ratio. Due to this, the adapted, transformed entropy optimization model with the empirical λ method returned the most risk-efficient portfolio.

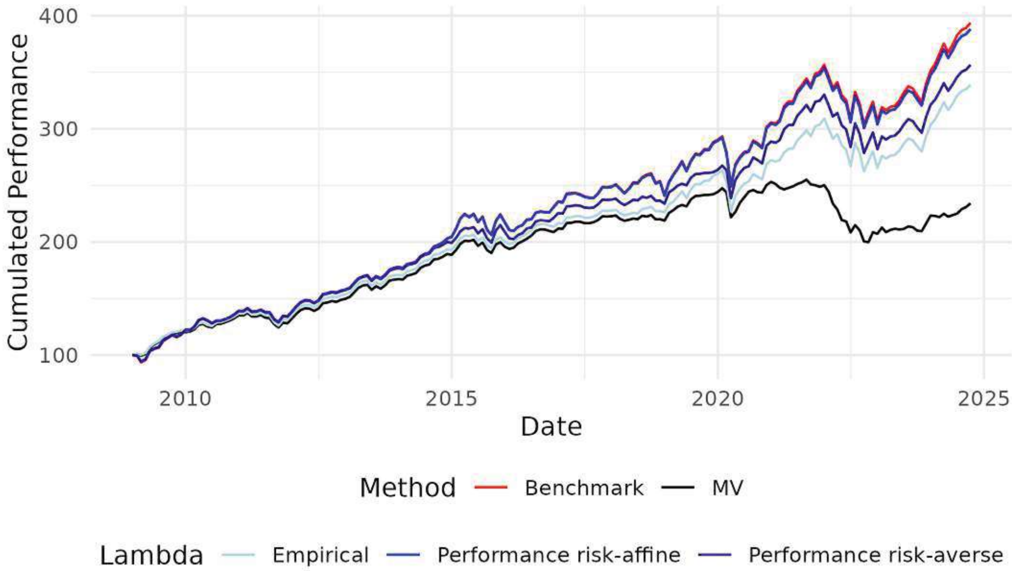

Figure 7.16: The performance of the adapted entropy model with entropy transformation for different λ calculation methods

8 Conclusio

All in all, the mean-variance method developed by Markowitz is a fundamental component of portfolio optimization regardless of its limitations and weaknesses. There are countless approaches to sophisticate and improve it and the inclusion of entropy works quite well. Both the entropy and adapted entropy method described in this thesis can deliver a more risk-efficient and diversified portfolio. The concentration in a few assets is the most focal point that is improved by these two models. However, one has to be careful with the different inputs and estimators used. A different way of estimating the expected returns of the assets, can change the weights of the whole portfolio.

Moreover, the model itself is also very sensitive to the parameter λ. The different ways of calculation always have a trend which direction the parameter tend to approach. Therefore, there might be more improvement, if a more robust way of calculating it is introduced.

Nonetheless, the goal of producing a more stable portfolio compared to the mean-variance method was succesfully reached. The portfolios calculated by the proposed models tended to have higher returns in the backtests as well as a better Sharpe ratio.

One of the questions of this thesis was also how to improve a certain benchmark. This was thoroughly tested, however not so easily answered. The problem of beating the performance of such a specified static portfolio in reality, is that it is strongly dependent on what the weights and the aim of said benchmark is. Risk-optimizing methods are improving the risk-efficiency of such benchmarks, however they tend to reduce the weight of volatile assets. However, they also come with the highest returns, resulting in weaker performance return-wise.

The inclusion of such portfolio optimization methods and understanding them is a crucial part of investing for several reasons. Utilizing mathematical and statistical techniques provides a solid foundation for investment decisions. They reduce the reliance on intuition and emotion biases, that can revolve in negative impacts on the performance. By systematically managing both the risk and the return, the growth of a portfolio can be controlled more consistently and sustainably contributing to a healty financial situation.

Bibliography

[1] Harry Markowitz. “Portfolio Selection”. In: *The Journal of Finance* 7.1 (1952), pp. 77–91.

[2] Ruidi Song and Yue Chan. “A New Adaptive Entropy Portfolio Selection Model”. In: *Entropy* 22.9 (2020). ISSN: 1099-4300. DOI: 10.3390/e22090951. URL: https://www.mdpi.com/1099-4300/22/9/951.

[3] Peter Joseph Mercurio, Yuehua Wu, and Hong Xie. “An Entropy-Based Approach to Portfolio Optimization”. In: *Entropy* 22.3 (2020). ISSN: 1099-4300. DOI: 10.3390/e22030332. URL: https://www.mdpi.com/1099-4300/22/3/332.

[4] “Normal (Gaussian) Distribution”. In: *Statistical Distributions*. John Wiley & Sons, Ltd, 2010. Chap. 33, pp. 143–148. ISBN: 9780470627242. DOI: https://doi.org/10.1002/9780470627242.ch33. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470627242.ch33. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470627242.ch33.

[5] “Lognormal Distribution”. In: *Statistical Distributions*. John Wiley & Sons, Ltd, 2010. Chap. 29, pp. 131–134. ISBN: 9780470627242. DOI: https://doi.org/10.1002/9780470627242.ch29. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470627242.ch29. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470627242.ch29.

[6] “Student’s t Distribution”. In: *Statistical Distributions*. John Wiley & Sons, Ltd, 2010. Chap. 42, pp. 183–186. ISBN: 9780470627242. DOI: https://doi.org/10.1002/9780470627242.ch42. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470627242.ch42. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470627242.ch42.

[7] Zvi Bodie, Alex Kane, and Alan J Marcus. *Investments*. eng. 6. ed. The McGraw-Hill / Irwin series in finance, insurance, and real estate. Boston, Mass. [u.a.]: McGraw-Hill, 2005. ISBN: 0071238204.

[8] Investopedia. *What Beta Means When Considering a Stock’s Risk*. 2023. URL: https://www.investopedia.com/terms/i/irr.asp#:~:text=The%20internal%20rate%20of%20return%20%28IRR%29%20is%20a,rely%20on%20the%20same%20formula%20as%20NPV%20does. (visited on 10/25/2023).

[9] corporatefinanceinstitute.com. *Yield to Maturity*. 2024. URL: https://corporatefinanceinstitute.com/resources/fixed-income/yield-to-maturity-ytm/ (visited on 10/12/2024).

[10] tradistats.com. *The Golden Cross*. 2024. URL: https://tradistats.com/golden-cross-an-der-boerse/ (visited on 10/12/2024).

[11] investopedia.com. *Earnings Yield.* 2024. URL: https://www.investopedia.com/terms/e/earningsyield.asp#:~:text=Earnings%20yield%20is%20the%2012,may%20indicate%20an%20undervalued%20stock. (visited on 10/12/2024).

[12] Martin J.Gruber Edwin J. Elton. *Modern Portfolio Theory and Investment Analysis.* 5th ed. Statistics in Practice. John Wiley & Sons, Inc., 1995. ISBN: 0471007439.

[13] Bernhard Pfaff. *Financial Risk Modelling and Portfolio Optimization with R.* Statistics in Practice. John Wiley & Sons, Ltd., 2013. ISBN: 9780470978702.

[14] Investopedia. *Beta: Definition, Calculation, and Explanation for Investors.* 2023. URL: https://www.investopedia.com/terms/b/beta.asp (visited on 11/30/2023).

[15] Investopedia. *A Deeper Look at Alpha.* 2023. URL: https://www.investopedia.com/articles/financial-theory/08/deeper-look-at-alpha.asp#:~:text=Alpha%20refers%20to%20excess%20returns,intended%20to%20eliminate%20unsystematic%20risk. (visited on 11/30/2023).

[16] Morningstar. *Should You Bank on Alpha?* 2023. URL: https://www.morningstar.com/articles/381488/should-you-bank-on-alpha (visited on 12/07/2023).

[17] inyova.ch. 2023. URL: https://inyova.ch/wp-content/uploads/2022/09/content_the-efficient-frontier.jpg (visited on 12/07/2023).

[18] Maciej J. Capiński and Ekkehard Kopp. "Frontmatter". In: *Portfolio Theory and Risk Management.* Mastering Mathematical Finance. Cambridge University Press, 2014, pp. i–vi.

[19] researchgate.com. 2023. URL: https://www.researchgate.net/profile/Federico-Guerrero/publication/269095610/figure/fig1/AS:462866848456704@1487367464155/Capital-market-line-diagram.png (visited on 12/15/2023).

[20] William F. Sharpe. "The Sharpe Ratio". eng. In: *Journal of portfolio management* 21.1 (1994), pp. 49–58. ISSN: 0095-4918.

[21] tickeron.com. 2023. URL: https://www.google.com/url?sa=i&url=https%3A%2F%2Ftickeron.com%2Ftrading-investing-101%2Fwhat-security-market-line%2F&psig=AOvVaw303iLxwgIXUGC_0KhSE7dA&ust=1702912907346000&source=images&cd=vfe&opi=89978449&ved=0CBEQjRxqFwoTCLiMz5HjloMDFQAAAAAdAAAAABAd (visited on 12/17/2023).

[22] Jean Dezert and Frédéric Dambreville. "Cross-Entropy and Relative Entropy of Basic Belief Assignments". In: *2023 26th International Conference on Information Fusion (FUSION).* 2023, pp. 1–8. DOI: 10.23919/FUSION52260.2023.10224207.

List of Figures