Mutual Fund Ratings – a Critical Analysis: Are Mutual Fund Ratings a Valuable Service for Investors?

Copenhagen Business School 2012

Cand.Merc. Finance and Strategic Management - Master Thesis

Author:	Kristina Roider
Supervisor:	Thomas Einfeldt
Pages:	79
Characters:	157,525
Submission Date:	14.11.2012

Executive Summary

Looking at today's investment market, we find an increasing number and complexity of mutual funds. Helping investors to cope with the investment decision, the mutual fund rating industry blossomed. Mutual fund ratings' influence is so high that researchers claim the ratings have impact on the in- and outflows and even the market share of the asset management industry.

Based on the rise and the high influence of mutual fund ratings, this study examines the value of mutual fund ratings for individual investors in Germany. The analysis is based on several different aspects of value. As ratings have a high influence on the fund flows it is necessary to ask whether fund ratings have predictive power for future fund returns. However, some raters claim that their rating should rather be used as a performance achievement score and only offers information value of past fund features for the investor. In order to answer both aspects, the analysis is divided into two parts.

From the first part of the analysis is found that the most important raters in Germany, Morningstar, Lipper and Feri employ different rating methodologies. Several components of mutual fund ratings that influence the quality of mutual fund ratings are outlined, before the three raters get evaluated based on these components. Each rating provider shows advantages and limitations in its ratings that have an influence on the quality assessment of the fund and therefore on the offered value for the individual investor. Detailed information about the fund is hidden from the investor in the complex rating methodology that produces one single rating score. This assessment of fund quality in one number communicates a relationship between rating and fund performance in the same way as credit ratings. This link therefore had to be analyzed in the empirical part of this study.

The empirical part of this thesis therefore tests the predictive power of mutual fund ratings from Morningstar and Feri for future return for the German fund universe from 2000 to 2011. This analysis is based on the most important study by Blake and Morey (2000). Results on this study indicate little predictive power for the Morningstar rating and no predictive power for the Feri rating.

Fund ratings offer little or no predictive power for future return in the mutual fund ratings and the complex process of fund evaluation also limits the offered information to the investor. The conclusion of this analysis is therefore that mutual fund ratings offer limited value to the individual investor in Germany.

Table of Content

1   Introduction

In the current mutual fund market we find an increasing number of funds with a vast variety of different investment guidelines and strategies (ICI, 2012a). The worldwide number of funds already exceeds the worldwide number of stocks traded on exchanges (World Federation of Exchanges, 2012)[1]. Besides the increasing quantity there is also an increasing complexity in the structure of funds. Investors who are willing to invest in the fund market find it therefore more and more challenging to process all available information and pick the most suitable fund for their investment portfolio.

With the main purpose of guiding investors through the fund investment decision the mutual fund rating industry blossomed. While the credit rating agencies had a bad credit crisis after 2007, fund raters gained more and more importance (Smith, Walter and DeLong, 2012). In order to help investors with the investment decision, first mutual fund ratings started with a pure comparison of past mutual fund returns. However, in recent decades there has been a development from this pure comparison of past returns to the practice of "fund screening" (Wittrock, 2003). This "fund screening" means that ratings process a range of measures that describe the characteristics of a fund besides its past return profile. In order to do so the agencies use different, sometimes complex rating methodologies in order to evaluate these attributes of the funds' past performance. For most rating agencies this assessment is based on past quantitative criteria, in some cases however also forward looking, qualitative indicators will be included in the evaluation process.

1.1   Field of Study

Several studies acknowledged this rise of mutual fund ratings and examined the high influence of fund ratings on the market. These studies for example find that the Morningstar rating, the most prominent mutual fund rating, highly influenced the fund in- and outflows (Del Guercio and Tkac, 2008) as well as the market share in the asset management industry (Khorana and Servaes, 2012). This phenomenon indicates that investors understand the mutual fund ratings as a predictor for future performance.

---

[1] According to the World Federation of Exchanges (2012) and the ICI (2011) there were a total of 45,508 listed companies and 69,519 available mutual funds worldwide in 2010.

However, Morningstar and other raters state that their ratings are not a guarantee for future performance and rather offer the investor a screening of the fund's general quality criteria of the past. The idea of this study is therefore to examine in a comprehensive manner what value mutual fund ratings really offer to the individual investors. Following this idea will help to shed further light into the field of mutual fund ratings. As research already explored the US market, the German market and its most prominent raters remained mostly unexplored. In order to close this gap and because Germany offers interesting insights into the European market, the focus of this study will lie on the German mutual fund market and its most important mutual fund raters. The problem statements will therefore be answered in the light of the German market.

1.2 Problem Statement

The question that arises from the rise of the mutual fund rating industry leads us to the question:

How valuable are mutual fund ratings for an individual investor?

As indicated above, the value of a rating can stem from different sources. Therefore, several sub-problems need to be clarified in order to answer the main question.

First and most intuitive is the predictive power of mutual fund ratings. As previous research showed investors might value the predictive power of a fund rating for future fund returns. A sub-problem to answer the question of the value of a fund rating is therefore:

- *Are mutual fund ratings a predictor for future fund performance in the German market?*

The value of fund ratings might however not be limited to the predictability of future returns. Most mutual fund raters state that their ratings are not a guarantee for future performance but can be used for a screening of different funds and their specific fund attributes. The use of different information aspects that are considered to assess the quality of the funds can therefore also contain value for the investor. It is therefore necessary to outline which fund information the industry covers in their ratings and

how they process it in order to offer it to the individual investor. Further sub-problems to answer the question of value in fund ratings can therefore be stated as follows:

- *How do the raters process and present the available fund information in the ratings and what are the differences between the different fund rating agencies?*
- *Where do the offered ratings show weaknesses and strengths in helping the investor to assess funds' past quality (performance)?*
- *What is the information for the investor received by the fund rating?*

In order to solve the outlined main problem it is important to assess all aspects of the outlined sub-problems.

1.3 Methodology

The following section describes the methodology of this thesis, namely the chosen structure and methods used when answering the problem statement. Furthermore it will be outlined which consequences these choices have for the conclusions drawn.

In order to answer the above stated problems several steps are necessary.

After describing the motivation and the brief methodology of this thesis, there will be a detailed overview over the literature on the topic of mutual fund ratings. In order to provide a better understanding of mutual fund ratings, a short comparison of mutual fund ratings and credit ratings follows, which will lead to the analysis.

In the first step of the analysis we will elaborate the different rating methodologies of the most important mutual fund rating agencies in Germany. The methodologies can generally be divided into the three steps of performance evaluation, fund categories and overall rating. After outlining these steps of the ratings, different secondary data from the mutual fund research will be used in order to describe the rating components and their importance for assessing the past performance of funds. The most important rating methodologies in Germany are then evaluated based on these elaborated components. The limitations and benefits of the different raters will be outlined in this part and compared to each other. The conclusions of this part of the thesis will give an answer to the last three sub-problems. As the evaluation of mutual fund ratings remained rather unexplored in the past, sources to answer these sub questions will mainly be secondary data from the research field of mutual fund performance.

In order to put more emphasizes on the predictive power of mutual fund ratings in Germany, the last part of the thesis attempts to answer the question if a fund rating is also a predictor for future performance in the German market. In order to do so an empirical study for the German mutual fund market will be conducted. To my knowledge this study will be the first to compare the actual predictive power of the mutual fund raters Feri and Morningstar for the German market within an identical examination period. In line with previous research a cross-sectional dummy variable regression will be used to investigate the predictive power of these raters for equity funds with a German investment focus. The precise regression methodology and data background will be elaborated in detail in the empirical part of this study. Primary data will be used to answer this sub problem.

In the Conclusion of this study the findings on all sub problems will be merged in order to answer the research question, if mutual fund ratings are valuable for the investor.

The thesis uses two different approaches in order to shed more light into the research problem. The first approach can be seen as a theoretical approach. The research methods of different raters will be explained and differences will be discussed in detail. Ground for this is the provided information material of the mutual fund rating agencies. As this study focuses on the German mutual fund market, the European Methodologies of the raters will be used. Using different findings from secondary data on mutual funds will then help us to find aspects that can determine the quality of a fund rating, as they influence the relative performance of funds. Based on these findings the fund ratings will be assessed and compared to each other in a comprehensive manner. By choosing this approach, the academic foundation of this part does not lie on a deep theoretical framework, but is rather driven by a practical application of the ratings.

The second approach is a pure empirical approach, where primary data will be processed in order to explore the predictability aspect of mutual fund ratings for the German market. The methodology of this empirical approach and the specific data will be explained in detail in the corresponding part of the analysis.

It is important to note, that this thesis makes use of secondary data in order to answer the stated research question. Secondary data has however always the draw-

back that it might originally be constructed for another purpose. However, the used secondary data of this study was selected carefully and where quality standards were in doubt this was clearly stated throughout the analyses and not used to draw conclusions.

Primary data used in this thesis is limited to the fund ratings and fund return data from the raters Morningstar and Feri. The data is of quantitative kind and the sources of this data can be considered trustfully, as Morningstar's performance data is regularly used in empirical studies.

1.4 Scope and Limitations

This study will focus on the question whether fund ratings are valuable for an investor in Germany. In order to narrow this research question several limitations had to be made.

Firstly, there is a variety of names for pooled investment vehicles that invest in securities, such as "investment trusts, investment companies, investment funds, funds, closed-end funds, open-end funds, mutual funds, unit investment trusts, fixed trusts, exchange-traded funds (ETFs), and hedge funds." (Fink, 2011, p.4). This thesis will only focus on the topic of mutual funds. In the course of this work these are defined as an investment vehicle of pooled funds that are invested in securities such as equities, bonds, currencies and similar assets. This is done by selling shares to the investor and investing "the proceeds in a diversified pool of securities, which are jointly owned by the funds' investors" (Anderson and Ahmed, 2005, p. 1). The key aspect is that mutual funds are subject to active portfolio management. Exchange-traded funds that are characterized by passive portfolio construction are not subject to this study. Furthermore, mutual funds are considered as open-end funds. That means funds are subject to redemption of the investor, i.e. investors can sell or buy shares based on the current value of the portfolio (Current net asset value, NAV) at any time (Fink, 2011).

Secondly, in the course of this thesis, we investigate the value of mutual fund ratings. The investigation of value is hereby not based on quantitative economic value theory but rather lies in the information theory. As ratings comprise different criteria of information and therefore build a source of information for the individual investor, the

more general research on information search is also important for the topic of mutual fund ratings. Following the pioneer work of Akerlof (1970) on information asymmetries in product information, research focused on the costs that arise from obtaining and verifying information (Stiglitz, 2000). In this course Barzel (1982) established a market framework that arises from measurement problems of product characteristics. Gerrans (2004) argues that this framework can be used as a proof for positive costs associated with the measurement of mutual fund quality. A mutual fund rating can therefore lower these costs, which is assumed to generate value for the individual investor.

Thirdly, as there developed a broad range of mutual fund rating agencies in the past years, only mutual fund rating providers, which are most influential in Germany, will be addressed. The influence of mutual fund ratings was decided based on the regular use of their ratings in information material by the four biggest German Asset Manager. Other providers will be omitted from the study due to time and space limitations.

Fourthly, the function of fund ratings is twofold. On the one hand, the rating is an assessment of a fund's risk-adjusted performance and its investment structure. It is used to inform the investor about the fund and its characteristics and can therefore be described as an information mechanism (Barzel, 1982 and Gerrans, 2004). However, this is not the only mechanism an investment rating shows. On the other hand persistent and independent performance evaluation of a fund can also be considered a control mechanism for the behavior of the portfolio manager (Eisenhardt, 1985). This study however does not focus on the management control mechanisms a fund rating has and how it affects portfolio manager's investment choice and risk taking. The paper solely targets the information aspect of the rating and its value for the investor.

Fifthly, the empirical part of this analysis only investigates the predictive power for Equity funds with a German investment focus that are for sale in Germany. This is in line with previous research and shows improvement to the study of the German market by Füss, Hille, Rindler, Schmidt and Schmidt (2010), who investigate all German mutual funds without any particular focus group.

Sixthly, there are several different approaches in previous research to address the predictive power and persistence of mutual fund ratings, like the dummy variable regression, the Spearman-rho rank correlation test or the use of the Marcov model. The cross-sectional dummy regression has however been the most popular and widely used one among scholar. In order to ensure the comparability of this study to other findings, the dummy regression was therefore used.

Seventhly, the study covers important information about Lipper. However, the predictability of the Lipper rating will not be covered in the empirical part of this study. This has several reasons. Firstly, Lipper does not provide a comprehensive rating for each fund, but five different ratings with different information variables. Using the dummy analysis from previous studies would therefore not be feasible in the same way. Using only one rating for the regression would however misuse the Lipper rating, as Lipper clearly indicates that the ratings should be used in conjunction with each other. Secondly, data on historic Lipper ratings is generally not available for research. Due to these limitations Lipper's rating will not be analyzed in the empirical part of this study.

Lastly, this thesis only focuses on quantitative mutual fund ratings and no qualitative ratings, as the Analyst Rating of Morningstar or Feri's rating for funds aged below three years will be subject to this thesis.

2 Literature Overview

With the rise of the mutual fund rating industry also academics developed interest in the topic of fund ratings. The research focus has however been quite narrow. Most studies focus on the U.S. market and the Morningstar Fund Rating. The following section gives an overview over the literature on mutual fund ratings.

The research on mutual fund ratings has primarily addressed three areas:

(1) The construction of mutual fund ratings with research from Blume (1998) and Sharpe (1998).

(2) The use of mutual fund ratings for individual investors with research from Capon, Fizsimmons and Prince (1996), Morey (2000), Jones and Smythe (2003) and Gerrans (2004). Including the question to what extend investors understand mutual fund ratings as a predictor for future performance with research from Damato (1996), Del Guerico and Tkac (2008) and Khorana and Servaes (2012).

(3) The predictive power of mutual fund ratings, which can be divided into three streams:

  i. The first research stream investigate the predictability of future performance with the use of mutual fund ratings with research from Khorana and Nelling (1998), Blake and Morey (2000), Gerrans (2006), Morey and Gottesmann (2006), Kräussl and Sandelowsky (2007) and Füss et al. (2010).

  ii. The second research stream investigated the performance persistence of mutual fund ratings with research from Duret, Hereil, Mitaine, Moussavi, and Roncalli (2008) and Hereil, Mitaine, Moussavi and Roncalli (2010).

  iii. The third research stream covers inhouse-studies from the mutual fund raters with research from Morningstar (2005), Morningstar (2005a) and Feri EuroRating Services AG (2012).

An overview of the different research areas of mutual fund ratings can be found in figure 1.
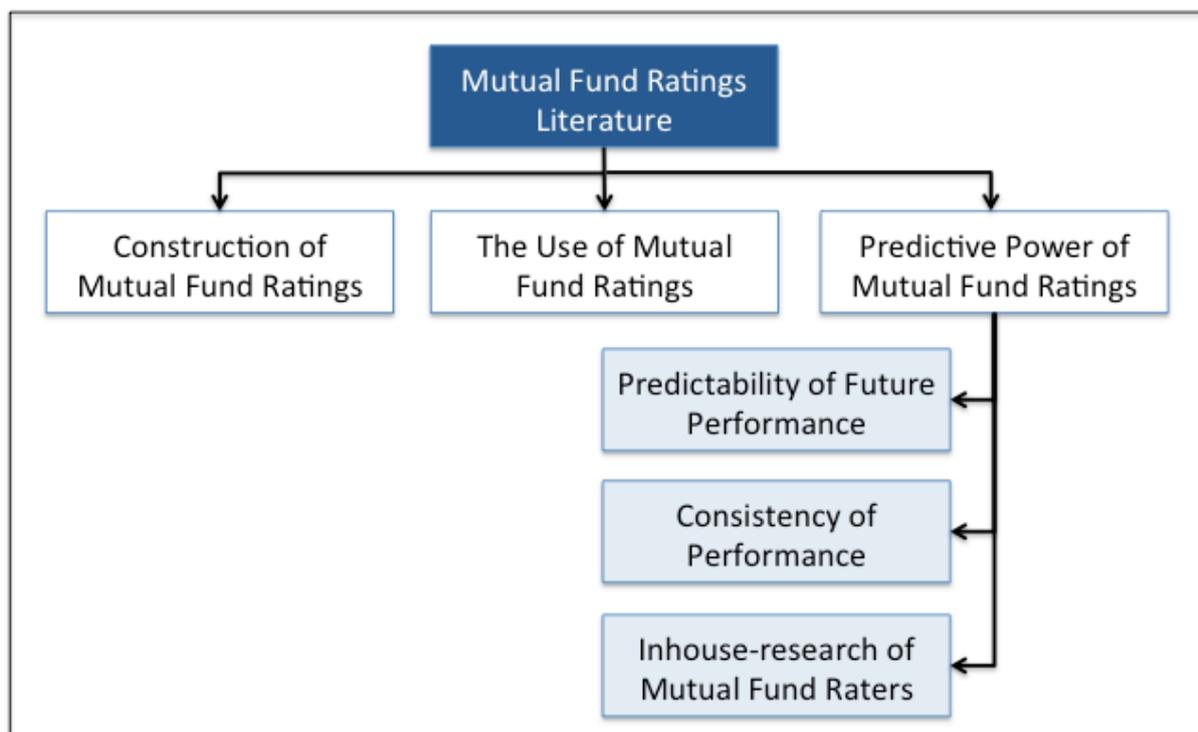
Figure 1: Literature Overview

Construction of Mutual Fund Ratings

The first research stream, the academic work on the composition of mutual fund ratings was solely focused on the Morningstar fund rating in the U.S. market and was mainly conducted in the 90's. Even though the research of Shape and Blume offers some interesting insights, the results are somewhat outdated since Morningstar changed its rating method in 2002 and 2007 in terms of their peer group categories and their risk assessment (Morningstar, 2009).

Use and Understanding of Mutual Fund Ratings for Investors

Research on the second stream, the mutual fund investment decision of individual investors shed light into the importance of mutual fund ratings. Capon et al. (1996) tested which variables influence investment choice for U.S. mutual fund investors. In a survey among mutual fund investors they found that historic performance rankings were the most important source used by investors followed by fund advertising. These findings are in line with Gerrans's study of the Australian fund market (2004), as Australian investors stated that ratings were the most influential input factor for the mutual fund decision process. An additional finding was that investors see the main purpose of ratings in the identification of the best-managed and administrated funds.

However, a very important finding of Gerrans (2004) is that "information source and selection criteria constructs are more useful in explaining the role of ratings than expectations of risk and return" p.87. Certain distinct investor groups put more emphasize on the fund criteria that build the input of the ratings than on the expectations for risk and return.

Individual investor's perception of mutual fund ratings has been in the focus of several more investigations. Researchers investigated the fund inflows according to the fund's rating change. Del Guercio and Tkac (2008) analyzed 10,000 Morningstar fund rating changes in the U.S. and reported that there are abnormal fund flows after a change in the star rating. These abnormal fund flows could not be explained by a change in other performance measures as the Sharpe Ratio or Jensen's alpha. The researchers therefore argue that the Morningstar rating has an independent influence on mutual fund in- and outflows. This is in line with an early study by Damato (1996), which was published in the Wall Street Journal that showed that in 1995 90% of new money inflows were invested in funds with a 4- or 5-star rating, while funds with a rating below 3-stars incurred a net outflow during the same investigation period. Khorana and Servaes (2012) investigate the influence of Morningstar ratings on the market share in the mutual fund market. By regressing the mutual fund market share by the Morningstar rating from 1992 onwards they find that the coefficient of the rating as independent factor is highly significant in explaining the market share of the mutual fund industry. This shows that Morningstar ratings do not only have a high influence on fund flows but also a considerable impact on the shape of the whole mutual fund industry.

Moreover, there is a number of studies which show that asset management firms are well aware of the influence the fund ratings have on investor's investment choice and use them as a signal to attract customers. A study from Jones and Smythe (2003) finds that 47.1% of the fund advertisements in *Money* magazine in 1999 contained performance evaluation from an independent fund research service, i.e. a mutual fund rater. Furthermore, Morey (2000) also argues that in the U.S. most advertisements from fund companies do not show any past performance measures besides their Morningstar Rating.

The third group is the most important stream for this work, as it covers mutual fund ratings' predictable power of future returns. The research question of predictability has been in the focus of a limited amount of papers.

The first study concerning the predictive power of mutual fund ratings can be attributed to Khorana and Nelling (1998). They used data on mutual fund performance and ratings from 1992 to 1995 and found that mutual fund ratings can predict persistence in the fund performance. The work of Khorana and Nelling has however been criticized widely. Among others Blake and Morey (2000) outlined that the results lack any credibility because of limitations in the used data and methodology. They argue that the examination period is too short and no adjustment for survivorship bias was made. Blake and Morey (2000) were then the first scholars to conduct a robust study in this area. The pioneers tested whether there is predictability of future performance for U.S. focus mutual funds in their Morningstar rating. The data set for this investigation contained out-of-sample performance data from 1993-1997 for all U.S. domestic stock equity funds rated in 1993. Furthermore in-sample return from 1993 to maximum 1983 was available, depending on the fund age. The fund ratings were taken from the beginning of the year. Out-of-sample performance measures of the study were the mean monthly excess return, the Sharpe ratio and the one- and four-index alpha. Blake and Morey (2000) then tested the relationship between these out-of-sample performance measures and the beginning-of-the-year fund rating for different horizons of one-, three-, and five-years.

They used a cross-sectional dummy regression and confirmed their findings with the Spearman-rho test. The cross-sectional dummy regression was in the form of:

$$S_i = C + \gamma_4 D_{i,4} + \gamma_3 D_{i,3} + \gamma_2 D_{i,2} + \gamma_1 D_{i,1} + \mu_i$$

Where S_i is the performance measure, C is the reference group of a five-star rating and $D_{i,1}$ through $D_{i,4}$ are the dummy variables that represent the star-ratings from one to four. Blake and Morey tested whether the different rating grades showed a distinct difference from each other in terms of fund performance. If the Morningstar rating shows predictability, the coefficients γ_1 through γ_4 should therefore be negative, as they measure the performance compared to the reference group, the five-star rating.

Furthermore the coefficients should also show a descending order $\gamma_4 > \gamma_3 > \gamma_2 > \gamma_1$, as there should be a distinction between the rating grades. In order to confirm the results from the dummy regression they tested the predictability with a non-parametric test of statistical dependence, the two-tailed Spearman-rho rank-correlation test. This method tests the correlation between ratings and fund returns. Furthermore, they tested how the Morningstar rating performs compared to other in-sample performance indicators calculated from the previous ten years, like the Sharpe ratio, Jensen's alpha, the four-index alpha or the historic mean monthly return in predicting future performance. The significance of these methods was tested with the t-statistics. With these two methods they discovered two robust findings. First, high-rated funds with a five- or four-star rating do generally not outperform 3-star funds. Second, low rated funds, i.e. funds with a one-or two-star rating perform significantly worse in the future than 3-star-or-above-rated funds. A further finding is that the Morningstar ratings are only slightly better in predicting performance than the other in-sample performance measures. The researchers therefore found proof for limited predictability of future returns in the US Morningstar rating.

As Morningstar's rating was subject to significant changes in 2002, Morey tested the new rating methodology in cooperation with Gottesman in 2006. They employed the same dummy regression analysis as in Blake and Morey's work in order to test the predictive power of the Morningstar rating for all domestic equity funds that were rated by June, 2002. Morey and Gottesman (2006) investigated the performance of these funds over the next three years until June 2005. Using the new Morningstar Methodology the results are quite different from Blake and Morey's paper (2000), as they find widespread support for the predictive power of the Morningstar rating within the three-year time frame. Especially, high rated funds outperform low-rated funds. Also the two-star funds outperform the lowest-rated funds, i.e. one-star funds significantly.

Several further studies were conducted in order to test Blake and Morey's findings in different country or asset settings. Gerrans (2006) for example tested the predictive power of Morningstar ratings for the Australian market. Based on Blake and Morey's dummy regression analysis he investigated the predictive power of the Australian Quantitative and Qualitative Morningstar rating for the category "Equity Trust" from 1996 to 2001. He used geometric monthly mean, the Sharpe ratio and Jensen's al-

pha as out-of-sample performance measures. He finds no consistent predictability for high rated funds for both ratings. However, in accordance with Blake and Morey's research he finds that the subsequent performance of a two-star rated fund is lower than for the five-rated fund. Kräussl and Sandelowsky (2007) test Blake and Morey's dummy variable regression for a longer time period from 1995 until 2005 and a larger sample of 25,202 funds from different categories of the U.S. market. They find that the Morningstar rating method is not able to beat random walk in the four broad asset categories but is able to distinguish good and bad performing funds within one single category. They also find that Morningstar's introduction of 64 fund categories in 2002 reduced the predictive performance of the rating system. Füss et al. (2010) test the predictability of the Morningstar rating for the German mutual fund market from 2004 – 2009 with the same dummy variable regression as the previous study from Blake and Morey (2000). They "could not reject the null hypothesis of no performance differences among five-, three-, four-, and two-star ratings in the majority of observation periods" (p. 85). However, they also found that the Morningstar rating has some predictability for worst rated funds, as this rating grade performed worse than the 5-star category in most observation periods. Füss et al. (2010) tested the predictability for the whole German fund universe and did not focus on funds with Domestic Equity focus like Blake and Morey (2000). Furthermore, the study was not adjusted for survivorship bias.

The second stream is characterized by work from Duret et al. (2008), who test the persistence of mutual fund ratings. They analyze the persistence of the five-star S&P and Morningstar Rating with the use of the Markov Chain. With this model they compute a transition matrix of the probabilities of a fund to remain in its five-star rating after a certain time. They find that there is no persistence for 5-rated funds. Hereil et al. (2010) also use the Markov Chain framework in order to investigate the rating persistence of Morningstar. Using the average investment period of individual investors, they judge the rating persistence as poor. They find that the Persistence Time, which is defined as the time period for which the probability of being downgraded is higher than the probability of remaining a 5-star rated fund is only 5 months for the Morningstar ratings. Compared to the Persistence Time of Credit Ratings, which is on average 10 years the mutual fund rating systems are much less robust.

A third research stream comprises of the inhouse research work conducted by the mutual fund raters. The mutual fund raters construct own studies on the predictability and persistence of their ratings in order to ensure and improve quality standards. For example, Morningstar evaluates in a paper the performance predictability of the U.S. star-rating after the methodology revision in 2002 (Morningstar, 2005). In a very simple approach they investigate the 3-year raw return of every rating grade from June 2002 and June 2003 and find that a five-star fund from 2002 shows a higher return than a lower-rated fund in some fund categories. However, there are also fund categories like International Equity where one-star rated funds outperform five-star rated funds in terms of raw return. Reason for the low rating might be the higher costs and risk, which are not captured by this simple return evaluation. Every year the Morningstar Research team investigates the performance of the star rating in this manner in order to assure its quality. These papers are however not available for every year.

Blake and Morey (2000) also report of an inhouse study by Laura Lallos from Morningstar in 1997 about the persistence of the Morningstar rating. She finds that 45% of all five-star funds in 1987 were still five-star-rated ten years later in 1997. However, the research does not contain any other results and very few details of the study were provided. Today, the paper is no longer available at Morningstar.

Morningstar conducted another study of the Morningstar Rating performance solely for the European market by the Morningstar Deutschland GmbH (Morningstar, 2005a). The evaluation period for this study ranges from 2003 – 2005 and includes all mutual funds of the European fund universe (around 25,000 funds). The study assumes that an investor picks funds in August 2003 based on the Morningstar star rating. Next, they evaluate the average performance and risk indicators over the next two years compared to the category mean and the star-grade (e.g. one-star group) mean. They find that the rating has a good predictability for selecting future high performing funds in the regional equity and bond categories. The use of the rating as a risk indicator works especially for the category of Small- and Mid-Caps. However, in the case of categories with a high risk level the rating only showed limited influence on the returns. Furthermore, the results were not tested for any statistical significance and are therefore of limited credibility.

Feri also tests the predictability of their rating (Feri EuroRating Services AG, 2012). Different from Morningstar, Feri examines the quality of its rating on a monthly basis. In the focus of their quality assessment lays the medium-term predictability. Since the rater started its monthly quality assessment in 1999, Feri claims that the performance of mutual funds with a A-, or B-rating significantly outperformed funds with a C-, D- or E-rating on average and that this is robust for the majority of fund categories. The outperformance of each rating grade is measured by its 3-year annualized return compared to the peer-group average. Particularly, for the fund category Equity Germany, Equity World and Equity Europe the outperformance of A and B rated funds was substantial. However, Feri performs no test for statistical significance. Besides the predictability, Feri also tests the persistence of its fund rating. Feri claims that 61% of all funds stay within the same rating grade over a year for the German fund universe. 21.3% experience a change by one rating grade and 2% experience a change by more than one rating grade.

3 Mutual Fund Industry

Before we take a closer look at mutual fund ratings it is helpful to outline the key characteristics of the mutual fund industry.

Since its first beginnings in 1924 the mutual fund industry has experienced considerable growth (Fink, 2011). The worldwide mutual fund industry achieved a total of more than 23 trillion USD in Total Net Assets in the last Quarter of 2011 (ICI, 2012). The highest growth rates in the mutual fund market appeared in the 90's (Fernando, Klapper, Sulla and Vittas, 2003) with the U.S. as the leader of this development. From 1992 to 1998 the total net assets in the USA increased from a USD 1.6 trillion to USD 5.5 trillion, implying an annual growth rate of 22.4% (Fernando et al., 2003). The members of the EU experienced similar growth magnitude with an increase in total net assets from USD 1 trillion in 1992 to USD 2.6 trillion in 1998, implying an annual growth rate of 17.7% (Fernando et al., 2003).

This increase in total net assets was accompanied by an increase in the number of mutual funds available around the world. In 2010 this number climbed to a total of 69,519 available funds around the world (ICI, 2011). In Europe alone the number of

funds increases every year by more than 1,000, with 3,400 new funds launched and 2,400 old funds closed (Moisson, 2012).[2]

These incredible growth rates were however not only fueled by institutional investments. A study from the ICI shows that the proportion of US household owning mutual funds increased rapidly over the same period. They report that this proportion grew from 27% in 1992 to 44% in 1998 (ICI, 2002). Mutual funds are therefore also a popular investment tool for individual and mostly unskilled investors.

4 Credit Rating vs. Fund Rating

Mutual fund ratings deliver an assessment of a fund's overall performance. The main purpose of credit rating agencies is to provide the individual investor with an assessment of the creditworthiness of entities and their obligations (Langohr and Langohr, 2009). Both ratings are therefore an information service for the investor. In order to outlay the special characteristics of the mutual fund rating industry we will take a closer look at the similarities and differences between the two types of ratings.

Long before mutual fund ratings evolved, credit raters became an integral part of the financial market place. The first credit ratings of the "Big Three", namely Standard and Poor's, Fitch Group and Moody's evolved at the beginning of the 20th century and became an integral part of the fixed-income market (Levich, Majnoni and Reinhart, 2002). The most important players on the market are currently the U.S. based companies Standard & Poor's, Moody's and the Fitch Group (Hill, 2002) with a market share of 97% in the U.S. (Bloomberg, 2011). Even though the rating agencies' reputation suffered during the US House Price Bubble, as many complex top-rated securities defaulted (Crouhy, Jarrow and Turnbull, 2008) the ratings have still a high influence in today's capital markets, as we can see in the developments during the Euro crisis.

Moody's explains in its investors guide (Moody's, 2012) that the credit ratings are based on a comprehensive analysis of quantitative ratios and on a fundamental analysis of the long-term prospects of the company. The process of assigning a rating is the same for all companies and governments. First, the raters gather information

---

[2] The study was conducted over the last 10 years.

from sources like annual reports, market and economic data, data from the meetings or conversations with the debt issuers but also from academic sources, central banks or ministries. Second, after analyzing this data the rating agency comes to a conclusion in a rating committee, where a diverse group of analysts with different specialty areas for industries, companies, asset classes and countries discusses the findings. Third, the development of the company or government is monitored during the time and might get reassessed if new information comes up.

Since the probability of default is always a forward-looking process, the ratings per se can only be on a subjective basis. Even though quantification is an integral part of every rating, the ratings are "not defined on a set of financial ratios or rigid computer models" (Moody's, 2012). This is also the standard of Standard and Poor's, which is outlined in Duret et al. (2008).

In order to get its fixed-income securities rated the entity has to pay a fee to the credit raters. Additionally, it has to offer the agencies insights into its financials and the raters usually schedule meetings and conversations with the representatives. The costs of a credit rating rose dramatically for the companies of Standard and Poor's, Moody's, and Fitch. Stemming from their market share of 99% for corporate issues and 98% for municipal bonds in the U.S. the three raters have a competitive position that allows them to set the market prices for a rating (Bloomberg, 2011). Moody's just raised its fee to 5 basis points of the amount being raised with a minimum of $73,000 (Bloomberg, 2011). S&P numbers are similar with a 4.95 basis point standard fee and $80,000 minimum (Bloomberg, 2011).

The standard approach for a fund rater is very different from the credit raters'. Most of the fund raters only take past performance of the fund into consideration. This past performance is however not only based on the return but also takes risks that stem from the investment and expense charges into account. Since fund ratings are quantitative there is no individual assessment of every fund and its managers on a case-by-case basis as it is the standard for credit ratings. The lower workload associated with a fund rating can surely be attributed to the fact that investment companies usually do not pay a fee to get their funds rated. However, the investment companies

pay a marketing license fee if they want to use the rating of mutual fund agencies in their information material or advertisements.

Unlike the credit rating industry the mutual fund rating industry is not as concentrated. The most influential fund rating agencies around the world include Morningstar and Lipper (DelGuercio and Tkac, 2008 and Herzog, 2007). Their ratings are not only of significance in the US but also around Europe and Asia. However, also companies we know for their credit ratings like Standard & Poor's, Moody's and the Fitch Group run a mutual fund rating office as part of their business. Additionally, there is a vast variety of domestic rating agencies that serve their domestic markets, like Feri for the European market, Euro Fondsnote for the German market and Aptimum for the French mutual fund market.

In the recent past there has developed a trend in the fund rating methods. Fund raters have announced that they want to offer investors a more forward-looking analysis. Morningstar for example introduced the Analyst Rating, which offers a forward-looking, subjective analysis of the biggest funds (Morningstar, 2012).

Table 1 gives an overview of the most definable differences between a credit rating and a fund rating.

	Credit Rating	Fund Rating
Time Perspective	Forward-Looking	Historical
Construction	Qualitative	Quantitative
Raters Influence	Subjective Rating	Objective Rating
Fees	Fees for rated company apply	Rating is generally free of charge
Conduction	Orderable at rating agency	Independent conduction based on investors' demand
Personal Contact	Personal contact with rated company obligatory	No interaction between rater and rated company required

Table 1: Comparison Credit Rating and Fund Rating

5 Mutual Fund Rating Agencies

As outlined above the mutual fund market is not as highly concentrated as the credit rating market. There are several important rating agencies that provide fund ratings

to the investor. As this study focuses on the German market it is important to elaborate which mutual fund ratings are most important in this market. When we take a closer look at the four biggest asset managers in Germany, namely DWS Investment, Allianz Global Investors, Union Investment and Deka (BVI, 2012), we find a pattern in their investor information material. All asset managers include the Morningstar and FERI rating of a fund in the Fund Factsheets that can be downloaded from the Internet and that is distributed to investors. However, some asset managers also include other ratings besides Morningstar and FERI, namely Lipper as can be seen in table 2.

Asset Manager	Ratings in Fact Sheets
DWS / Deutsche Bank	Morningstar, FERI, Lipper
Allianz Global Investors	Morningstar, FERI
Union Investment	Morningstar, FERI, Lipper
Deka	Morningstar, FERI

Table 2: Ratings included in the Factsheets for the four biggest German asset manager

As the three companies are the most dominant fund raters in the biggest asset managers of the German mutual fund market, the analysis will look into the rating methods of Morningstar, Lipper and FERI. This will help to give an answer to the sub problem on *How the raters process and present the available fund information in the ratings and what are the differences between the different fund rating agencies.*

5.1   Morningstar Inc.

[T]he brand that has emerged as dominant in the 1990s is not Fidelity, Putnam or even Merrill Lynch—but instead is Morningstar." R. Pozen, The Mutual Fund Business (1998), p. 75.

The most dominant and therefore most important mutual fund rating agency is Morningstar Inc. (Morey, 2002). Morningstar's fund rating methodology can be traced back until its start in 1985 (Morningstar, 2010). The Chicago-based company started with only 400 funds and now keeps a database of more than 375,000 investments, including shares, mutual funds and others (Morningstar, 2012a). The number of rated

funds for the German market already climbed to a total of 26,562 (Morningstar 2012b).

Its famous rating from one to five stars has already become integral part of the mutual fund vocabulary (Morey, 2000). Investors find the evaluation of a mutual fund in a single rating convenient and think its famous 1- to 5-star rating key is as easy to understand as the star-based rating of a hotel or of a restaurant (Morey, 2000).

It is therefore of special interest to evaluate the value of Morningstar's rating for investors and to outline the methods Morningstar uses in order to arrive at its famous star rating. Morningstar's rating is assigned from one to five stars based on the risk-adjusted return. This performance is then measured against its peer group (Duret et al., 2008) and the funds are rated accordingly. In order to analyze this rating method the three general steps will be outlined in more detail in the following subsections.[3]

5.1.1 Morningstar Performance Evaluation

In order to receive a fund rating the mutual fund has to be available on the market for more than three following years. Morningstar uses the past performance on these funds and adjusts it by the funds' risk properties and costs in order to arrive at the funds' Morningstar risk-adjusted return (MRAR).

The exact procedure consists of four steps. The first step is to calculate the Total Return (TR_t) based on the following formula:

$$TR_t = \left\{ \frac{P_e}{P_b} \prod_{i=1}^{n} \left(1 + \frac{D_i}{P_i}\right) \right\} - 1$$

Where P_e and P_b are defined as the NAV of a fund, the net asset value of the fund at the end and the beginning of the month t. D_i is the per share distribution, i.e. capital return, dividends or distributed capital gains. P_i is defined as the amount of net asset value per share that has been reinvested at time i. n is the cumulative number of distributions during month t. Underlying assumption is that investors do not pay any

---

[3] The description of the Morningstar fund rating method will be based on the Morningstar Methodology Handbook (2009) available at
http://corporate.morningstar.com/de/documents/MethodologyDocuments/MethodologyPapers/Morning starFundRating_Methodology.pdf if not indicated otherwise.

transaction costs and reinvest all their distributions. The percentage Total Return is therefore the increase/decrease in the NAV at the end of the month compared to the beginning of the month multiplied by the share of all distributions over one month of all NAV gains reinvested.

In the second step all fees and loads are taken into consideration and the return is adjusted for these obstacles. The Load-adjusted Return for month t is defined as:

$$LR_t = (1 + TR_t)(1 - F)(1 - R_t) - D_t(1 - F)\frac{\min(P_0, P_1)}{P_0} - 1$$

where F is defined as the maximum front load, R_t is the redemption, D_t is the deferred load, all as decimals for time t. P_0 is the NAV at the beginning of the period per share and P_1 is the NAV at the end of the period.

Third step is to calculate the Excess Return (ER_t) based on the current risk-free rate (RF_t):

$$ER_t = \frac{1 + LR_t}{1 + RF_t} - 1$$

For European Funds the risk-free rate is either the Merrill Lynch one-month LIBOR index or the three-month Treasury Bill, depending on the fund category.

The forth and last step takes the risk aversion of investors into consideration. This is done by equalizing the certainty equivalent of the excess return to the excess return adjusted by risk aversion:

$$E(ER^{CE}) = (E(1 + ER)^{-\gamma})^{-\frac{1}{\gamma}} \qquad \gamma > -1, \gamma \neq 0$$

where ER^{CE} is defined as the certainty equivalent excess return and γ is a measure of risk aversion. The Morningstar risk-adjusted return (MRAR) is then generated by replacing the expectation with its sample estimate:

$$MRAR_t = \left[\frac{1}{T}\sum_{i=1}^{T}(1 + ER_t)^{-\gamma}\right]^{\frac{12}{\gamma}} - 1$$

Morningstar uses a value of $\gamma = 2$ for the risk aversion of the investor. This implies

that investors' risk-aversion is a concave function of returns meaning that investors show a decreasing marginal utility with increasing returns. Morningstar (2009, p.14) comments that this is also taking the higher downside risk into account, as it is "steeper for negative returns and starts flattening out for positive returns, and this puts more emphasis on the downside variation."

Before 2002 Morningstar used a different method to adjust the fund performance for risk. A fund's risk was measured as the average underperformance against a 90-days Treasury bill. Funds that never fell behind this benchmark were classified as riskless. However, Morningstar adjusted the rating methods after the dot.com-bubble. The problem was that for the years before the collapse, technology funds never underperformed the risk-free rate benchmark and were therefore considered as risk-free. However, as the bubble burst in 1999 these funds incurred huge losses. Morningstar now employs a different risk adjustment approach in order to avoid such misleading ratings. It is based on the expected utility theory and takes two findings about the risk perception of investors into consideration:

(1) Investors are more concerned with unexpected losses than with unexpected gains.

(2) Investors are willing to give up some expected gain in order to increase return probabilities.

Morningstar changed its risk assessment in 2006 in Europe and now accounts for all variability in fund returns but puts greater emphasize on the negative outcomes.

5.1.2 Morningstar Fund Categories and The Morningstar Style Box™

After the risk-adjusted performance and the weighted rating have been calculated, the fund has to be compared against the competitors from its category peers. The assignment of a fund to a specific category is based on the underlying securities in a portfolio.

Morningstar distinguishes its funds according to the Morningstar Style Box™. This distinction of funds was first introduced in 1992. The reason why the fund rater employs this style criterion is investors' search for diversification (Morningstar, 2004). Diversification is not just a buzzword in the investment landscape. Investments that show a correlation coefficient below one to the existing investment portfolio reduce

the risk of the overall risk for the investor when added to the portfolio (Markowitz, 1952). Investors are therefore often searching for a specific asset class among mutual funds that decreases their overall portfolio risk. The Morningstar Style Box™ is helping them to differentiate the variety of fund styles and to build better portfolios.

Morningstar takes two measures to locate a fund in the style box. First the market capitalization of the fund's stock holdings, second the value-growth investment style of the fund. It distinguishes between large-cap, medium-cap and small-cap funds based on the fund's stock holdings' market capitalization and between growth, blend and value funds based on the fund holdings' value and growth properties. The resulting nine categories are depicted in the box format shown in figure 2.
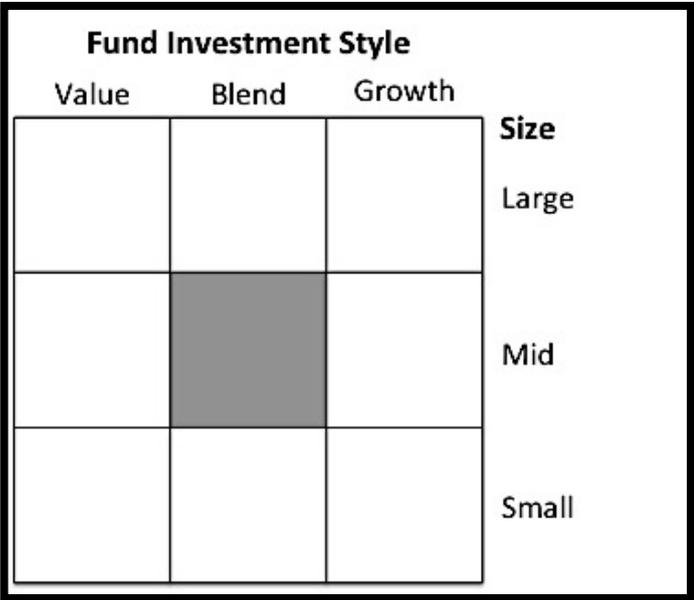
Figure 2: Morningstar Style Box (Morningstar, 2004)

However, Morningstar does not use the Morningstar Style Box in order to categorize funds to assign the Star rating. Morningstar employs much more detailed investment style categories, which can be seen in the Appendix 3.

For Europe the number of categories was increased in 2006 and settled for 142 Morningstar categories. The number rose consistently to more than 200 categories in 2012 (Morningstar, 2012b). Morningstar distinguishes four broad categories in Europe: Equity, Bonds, Balanced and Others. For European, US or other main investment focus-funds also the style criterion of the Morningstar style box is used to distinguish the high variety of funds in this segment.

5.1.3 Morningstar Final Rating

All funds of one investment category (peer group) are then ranked based on their MRAR-value described in 5.1.1. However, the funds' MRAR is not equally weighted over its existence. Morningstar adjusts the performance for the age of the fund using the metrics outlined in Table 3. The Overall Morningstar Rating of a fund is therefore a weighted average of its separate 3-year, 5-year and 10 year rating. Condition is however that the fund did not change its category in the addressed time frame.

Performance Record	Overall rating
36 – 59 months	100% three-year rating
60 – 119 months	60% five-year rating
	40% three-year rating
120 months or longer	50% 10-year rating
	30% five-year rating
	20% three-year rating

Table 3: Morningstar Rating Weights (Morningstar, 2009)

After weighing the fund performance measured by the MRAR-value and assigning it to a specific peer group the fund will be ranked among its peers. The 10 % with the highest weighted MRAR-values within an investment category receive a 5-star rating. The next 22.5% receive a 4-star rating. The following 35% receive a 3-star rating. The next 22.5% in the category receive a 2-star rating and the last 10% of the funds receive a 1-star rating. This rating distribution within one category is based on the characteristics of a bell curve as can be seen in figure 3.

The rating of existing funds will be reassessed on a monthly basis. The peer group classification of a fund will be reassessed every 6 months.
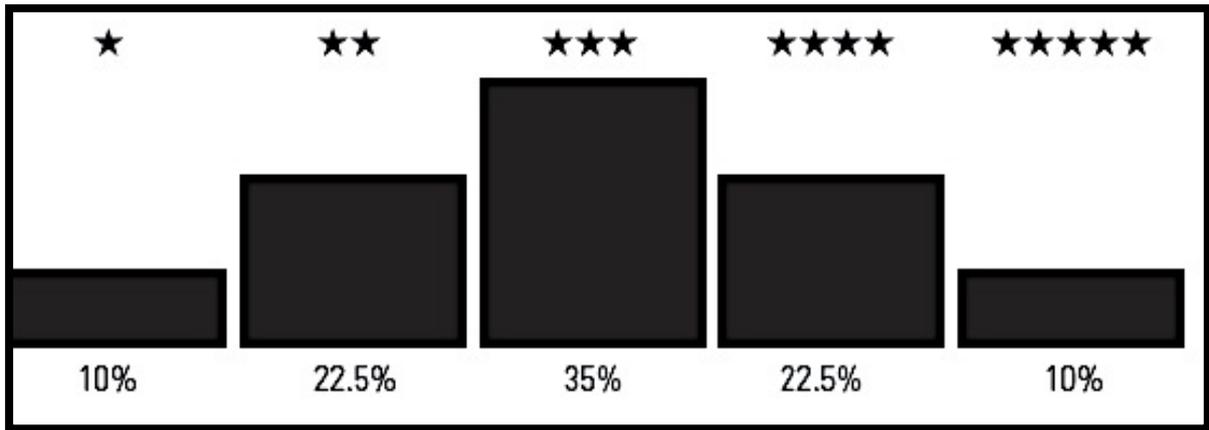
Figure 3: Morningstar Star-Rating Scale (Morningstar, 2009)

5.2 Lipper Rating

Lipper is a Thomson Reuters Company and provides information on mutual and hedge funds, commentaries, fund awards and rating information to individual investors and financial advisors. Its "Lipper Leader" fund rating is considered as one of the most important ratings besides Morningstar and comprises more than 117,000 rated funds in 61 countries. Even though the Lipper rating is not as popular as the Morningstar "Star Rating" it is still very common in fund information, and the Wall Street Journal as one of the most influential economic newspaper is publishing it along with its famous mutual funds Scorecards (Wall Street Journal Homepage, 2012).

The rating key shows some similarities to the Morningstar Star Rating. Lipper also differentiates between five different rating grades for the funds. A fund can achieve a rating from 1 to 5, where 5 is best grade also called "Lipper Leader" and 1 is the worst grade. The exact rating key and Lipper design is shown in figure 4. Lipper's three steps of performance evaluation, fund categorization and the final rating will be outlined in the following parts.

Figure 4: Lipper Leader Rating Key (Lipper 2012b)

5.2.1 Lipper Performance Evaluation

Lipper's self-stated main goal is to advice individual investors and financial advisors which funds are most appropriate for their investment style and their investment goals.[4] The criteria Lipper uses to give such advice are in Lipper's words 'investor-centered', i.e. concentrate on consistent, strong returns (Lipper, 2008).

An important difference to the Morningstar Rating is that funds are not just rated based on one metric, but are rated based on five key metrics, namely Total Return, Consistent Return, Preservation, Tax Efficiency and Expenses. That means that a fund can get up to five different ratings. The rating method for each metric are outlined in the following:

(1) **Lipper Ratings for Total Return:** For this rating metric the relative historical performance of a fund compared to its peers will be assessed. The performance is however not adjusted for risk. Total Return is just defined as the historic return net of expenses, taking reinvested dividends into account.

Investors who are only looking for possible high returns and are also willing to take the downside risk of an investment might be interested in this category. This rating metric is stated to be important for investors who are expecting a bull-market in a specific fund category.

(2) **Lipper Ratings for Consistent Return:** This rating metric compares the historic risk-adjusted returns of funds among a peer group. The rating considers

---

[4] The description of the Lipper Fund Rating will be based on the International Methodology Handbook (2012) available at http://www.lipperweb.com/docs/AboutUs/LLMethodology_Intl_V2_A4.pdf
and on the US Methodology Handbook (2011) available at
http://www.lipperweb.com/docs/research/leaders/LLMethodologyUS_V3.pdf, if not indicated otherwise.

long-term and short-term risk-adjusted performance with a measure for consistency. This measure is founded on the Hurst-Holder exponent and on the principle of Effective Return.

The H exponent is a mathematical tool that measures the deviation from the random walk. The exponent can achieve values from 0 to 1. It is designed to show a value of 0.5 in case of random walk in the data series. Random walk in this case means that the returns of the fund are not predictable and that stock price changes are just as likely to be high as low (Spritzer, 2001). If the value is above 0.5, the data shows positive autocorrelation and persistence, i.e. a move in one direction tends to be followed by a move in the same direction. E.g. a high return will be followed by another high return and the returns a long time into the future will also tend to be high. This means that returns are more persistent than a random walk (Stutzer, 2005). If the value is below 0.5, the data shows negative autocorrelation (mean reversion) and persistence, i.e. a move in one direction tends to be followed by a move in the opposite direction. E.g. a high return will probably be followed by a low return and the value after that will tend to be high. Persistence also indicates that this trend to switch between high and low values lasts a long time into the future. This means that the returns are less persistent than a random walk (Reiter, 2007).

The Hurst Exponent is therefore "a global measure of risk, defined as the smoothness or unsmoothness an asset exhibits." (Clark, 2003, p. 9). This is in line with Mandelbrot (1963) who defines the H exponent as the intrinsic volatility when volatility is defined as the smoothness of the sample path.

Since investors are interested in funds that show a smooth trend in their returns, i.e. a H value above 0.5, Lipper sorts the funds according to their H value into three categories. Funds with an H exponent below 0.45 will be assigned to the last category, since they show the lowest smoothness in the time series. Funds with an H exponent between 0.45 and 0.55 will be assigned to the medium category. The highest category is reserved for funds with an H exponent above 0.55, since they show the highest smoothness in the time series.

However, the H exponent is not the only measure that Lipper uses in order to define a fund's Return Consistency. A high H exponent can also indicate persistence and serial correlation for low returns, i.e. a relatively low return is likely to be followed by a relatively low return a long time into the future. It is therefore important to differentiate between loss and profit. This is where the Effective Return (ER) plays a major role. Effective Return is a risk-adjusted investment performance measure developed by Dacorogna, Gençay, Müller and Pictet (2001). It takes the risk aversion of the investor into account and allows for a lower risk aversion in the area of gains and for a higher risk aversion in the area of losses.

Lipper therefore investigates if funds with high H exponents also show positive Effective Return. Funds that have a high H value but a negative Effective Return will be assigned to the last category with H values below 0.45.

The rating process for the Return Consistency can be summarized as follows:

(1) All funds from one peer group are sorted by their H value in a descending order.

(2) The funds are assigned to three different groups based on their volatility behavior, measured by the H exponent. Funds with an H exponent above 0.55 will be assigned to the highest category. Funds with an H exponent between 0.45 and 0.55 will be assigned to the medium category and funds with an H exponent below 0.45 will be assigned to the last category.

(3) In each of the three H exponent groups funds are now reordered based on their Effective Return value, starting with the highest and proceeding to the lowest.

(4) The funds with a low or even negative ER value in the first group of highest H exponents (H above 0.55) are removed from this group and placed at the bottom of the last H exponent group (H below 0.45).

(5) The 20%-rating distribution is now applied to this fund ranking, assigning the values of 1 to 5 for Consistent Return Rating.

Risk-averse investors that are concerned about the downside risk of a fund will be interested in the Rating of Consistent Return. However, the investor has to be aware of the fact that some fund categories inherent more risk than others. Therefore investors should also be careful comparing Consistent Return Ratings from different categories with each other.

(3) **Lipper Ratings for Preservation:** This rating evaluates funds based on their loss characteristics. It intends to measure the historical loss aversion of a fund among its asset class, i.e. equity, mixed-asset or fixed income. In order to do so the measure shows a one-parameter estimate of a fund's downside risk. Lipper claims that this category is of special interest for the investor with absolute loss avoidance.

The measure is defined as the sum of negative monthly returns either over the 3-year, 5-year, 10-year or overall period. The benchmark for this measure is 0 in order to separate the positive and negative returns (Stutzer, 2005). The Lipper Rating for Preservation (LP) can be expressed by:

$$LP = \sum_{t=1}^{n} Min\,(0, R_t)$$

where n = 36, 60 or 120 months, and R_t is the return in month t (Amenc and Le Sourd, 2007 and Stutzer, 2005). The highest number that can be achieved by this Lipper Preservation Rating is 0.

This rating is the only Lipper Rating that does not evaluate the fund compared to its peer group. Funds compete with all funds within their asset class, i.e. equity, fixed-income and multi-asset.

However, the investors that rely on this measure should also be aware that some asset classes have historically been more volatile than others, for example equity funds have a higher volatility than multi-asset of fixed income funds.

(4) **Lipper Ratings for Tax Efficiency:** In this rating metric funds are ranked according to their tax efficiency measured by a fund's ability to postpone taxable

distributions compared to its peer group. This rating is only available for U.S. funds and will therefore not be in the focus of this research.

(5) **Lipper Ratings for Expense:** This rating compares a fund's expense ratios among its competitors with similar load structures. The rating is of importance since gross return will be diminished by higher expense associated with the fund. A fund with a high expense ratio therefore has to show better ratings in other categories in order to compensate the extra outlay of the investor.

Lipper distinguishes three different load-categories: no-load/front-end load, back-end load/ level load, and institutional load. It groups all funds within a peer group according to these three load categories and then ranks these funds based on the expense ratios for the 3-year, 5-year and 10-year period.

5.2.2 Lipper Fund Categories

As of 2008 Lipper used a total of 289 fund categories in Europe (Lipper, 2008). Lipper claims in the category guide (Lipper, 2008) to use a fund classification methodology that creates "homogeneous groups of funds with comparable investment objectives" (Lipper, 2008, p.2). An overview over the categories can be found in Appendix 4. Funds within one peer group have investments within the same financial markets, but are not obligated to have the same investment style or strategy. A category has to contain at least 10 funds in order to get rated. However, also funds from categories with less than 10 funds may receive a rating upon request.

Generally at least 75% of a portfolio have to be invested in line with the classification of the fund. However, also the long-term composition is taken into consideration, since short-term changes in strategic asset allocation are tolerated. Lipper uses the stated investment objective, the fund prospectus, the fund Fact Sheet and the annual and semi-Annual Reports in order to determine the fund classification. Main input factor is however the stated classification of the investment company.

Broad Fund Categories are Equity, Bond, Mixed Assets and Others. Equity funds are distinguished based on Region, Country and Industry. Funds that follow a small- or mid-cap strategy will be assigned accordingly. Funds in the Bond classification contain fixed-income products with average maturity above one year. The classification of the funds will be based on the currency exposure, emerging market exposure,

credit quality and maturity. Mixed Asset Funds are categorized by their risk level, country or region focus and currency.

5.2.3 Lipper Final Rating

Lipper employs the same final rating method for all five rating metrics after the funds were ranked according to their scores. Like the Morningstar rating, a fund's rating is based on its performance compared to its peer group. The distribution of the ratings among a peer group is however different from Morningstar's. For the Lipper Ratings the peer group is divided in 20%-steps. As figure 5 shows the 20% best funds in a peer group achieve a rating of 5, the "Lipper Leader" status. The following 20% are assigned a rating of 4. The next 20% are in the rating category of 3. The next 20% are allocated to the rating grade 2 and the last 20% to rating grade 1.
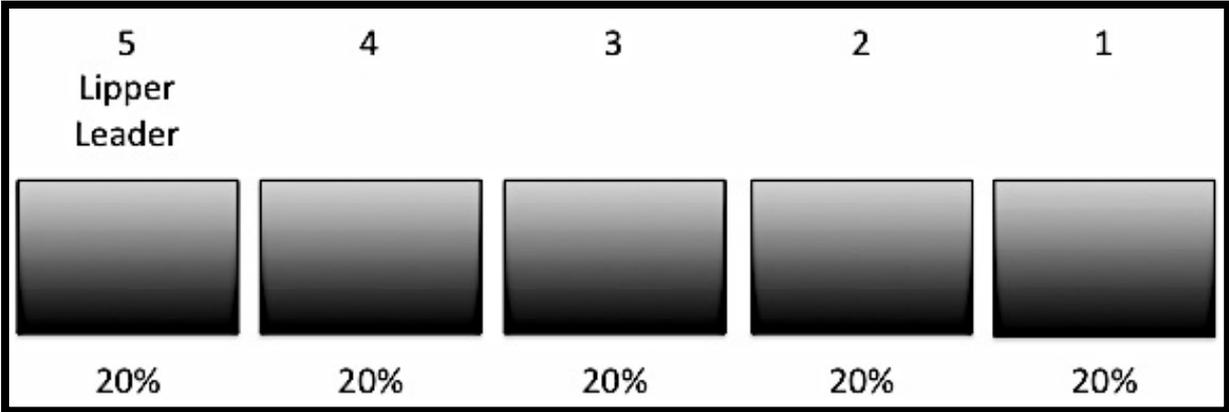
Figure 5: Lipper Rating Distribution (Lipper, 2012b)

Ratings are calculated for 3-years, 5-years, 10-years and overall periods, whereas the final rating is calculated as the equal weighted average of the percentiles over the four different periods. The same procedure as for the Morningstar revision applies for the Lipper rating, as fund ratings are also reviewed every month. However, fund classifications are updated on an ongoing basis. If new products come up, Lipper may also introduce new categories.

5.3 FERI Rating

FERI EuroRating Services AG (in the following referred to as Feri) is one of the few European-based fund raters. The company published its first ratings in 1992. The Feri AG is based in Germany but also developed considerable business activities in

other parts of Europe, especially in Austria, Switzerland, France and the U.K. (ft.com, 2012) but also in Italy and Sweden (Feri, 2012). Over the time Feri increased the number of rated funds rapidly. While it rated a total of 936 funds in 1999 the number rose to a total of 8.399 ranked funds in 2010.

Feri's aim is to create a rating that provides true support in the investment decision the investor faces on the current fund market.[5] Feri's belief is that a fund rating can only provide added value for the investor, if it is a predictor for future performance. Based on this perception, Feri states that its rating is fulfilling this necessary predictability of future performance and therefore is valuable for an investor. When compared with Morningstar and Lipper, Feri is the only fund rater that actually claims its ratings are a predictor for future performance.

In order to fulfill this predictability Feri argues that its Feri EuroRatings are mostly concerned with the quality of a fund. A fund's quality arises from the fund management's performance but also from the fund's ability to handle risks.

The Feri rating key contains five rating grades outlined in table 4. Where A and B ratings are supposed to classify funds that will show a stable, above average performance with relatively low risk over the long-run. A rating grade of C classifies funds that show average performance. Below average funds are assigned to class D and E.

A	Very good
B	Good
C	Average
D	Below average
E	Weak

Table 4: Feri Rating Classes

After a change of the portfolio management team the ratings are all labeled with the supplement 'ur' ('under review'). If the fund is not showing continuity in the next 12 or 18 months after the change, the fund rating will be withdrawn.

---

[5] The description of the Feri EuroRating method will be based on the Feri EuroRating Services AG (2012) Methodology available at description
http://www.feri.de/Content/Frr/Files/Fondsplattform/PUBLICARCHIVE/Fondsrating.pdf if not indicated otherwise.

5.3.1 FERI Performance Evaluation

The Feri rating team decides individually which funds to rate. However, the funds have to meet four conditions in order to be eligible for a rating:

- The fund has to show at least five years of fund history.
- The fund has to be available for public sale.
- There was no recent strategy or investment style change.
- The fund can be allocated to an investment style category with at least 20 funds, which fulfill the first three conditions.

The reasoning behind a 5-year fund history is based on the cyclical return behavior of mutual funds. Feri argues that five years build a good basis for the performance evaluation through different stages of the business cycle.

However, Feri recently changed the first condition and now also rates funds that record less than five years of past data on a fee basis. The construction method for these young funds is different. Because of the short evaluation period funds are not only rated based on their past data, but also on qualitative indicators, like questionnaires and interviews with the investment company. Feri only rates such young funds upon the investment company's request. The different weights of qualitative and quantitative indicators for different fund histories are illustrated in figure 6.
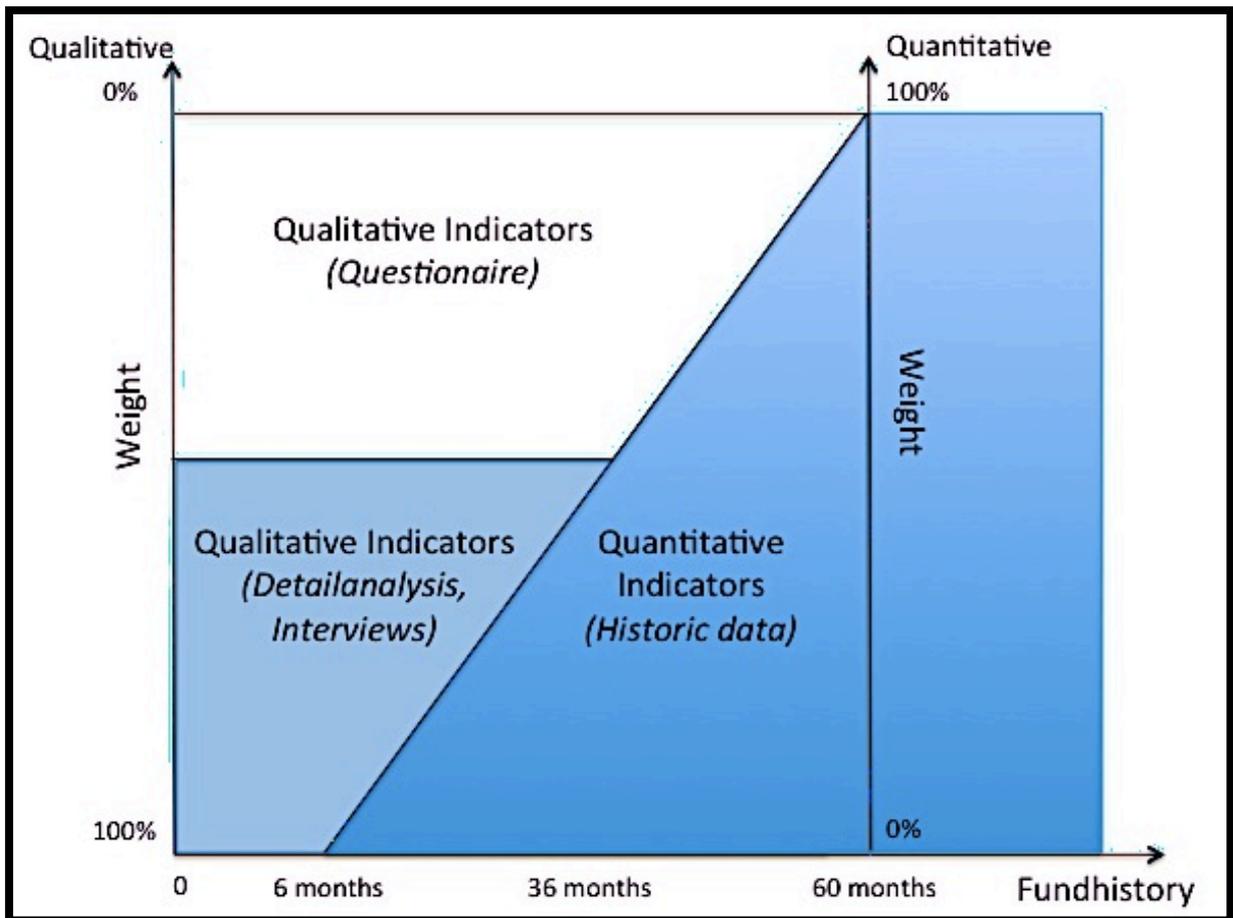
Figure 6: Weight of Qualitative and Quantitative Indicators

The weight that is put on the qualitative indicators depends on the age of the fund. Young funds consist of a much higher proportion of qualitative indicators than older funds. A fund, that has no or less than six months of historical data is solely evaluated by its qualitative indicators. A fund, which is older than five years, is solely evaluated by its quantitative indicators.

As outlined above Feri claims that the quality of a fund depends on performance and risk indicators. Feri therefore takes both into consideration when rating a fund. Table 5 indicates the weight of each category and its individual factors:

Performance: weight 70%	Risk: weight 30%
Relative Performance	Timing Risk
Long term Profitability	Risk of Loss
Stability	Behavioral Risk

Table 5: Structure of the Feri Rating Indicators

The Performance aspect of a fund will make up for 70% of the overall rating. Table 6 gives an overview over the different criteria determining the performance evaluation in course of the fund rating.

Criteria	Performance Indicators	Weight
Relative Performance	1. Outperformance of the Index p.a.	15%
	2. Average rank among peer group (3 months rolling)	20%
Long-term Profitability	1. Positive Elasticity	15%
	2. Difference in Elasticity (-/+)	10%
Stability	1. Probability of Outperformance against Index (3 months rolling)	20%
	2. Probability of Outperformance against Peer Group (3 months rolling)	20%

Table 6: Performance Indicators for Feri Fund Rating

The performance aspect of the fund will consider three areas: Relative Performance, Long-term Profitability and Stability.

In the assessment of a fund's relative performance the emphasize lies on a fund's potential to generate outperformance relative to the benchmark but also to the peer group. In the quantitative evaluation financial ratios give an indication if the fund is only displaying short-term outperformance compared to its peers or if the fund convinces with continuity in its performance.

In the evaluation of a fund's long-term profitability its positive elasticity plays a major role. Positive Elasticity means that a fund is able to capture upward movements of the market above average standards. This means Feri investigates how fund returns behave if the market is going up by 1%. Furthermore, Feri looks at the difference in elasticity, i.e. the difference of positive and negative elasticity. This is an indication of how the fund reacts to the different up- and downward movements in the market. E.g. if a fund's difference in elasticity is negative, it is losing more return in downward movements than it is gaining in upward movements of the market. This key figure therefore investigates the behavior of a fund when the benchmark is moving sideways. If the number is positive, the investor is winning money in a sideways movement.

The last performance criterion is the stability of a fund. In order to quantify this criterion, Feri evaluates if a fund is outperforming the index or its competitors during any 3-month period. The rater then compares the number of months the fund underperformed with the number of months the fund outperformed.

The remaining 30% of the Rating is based on three different risk criteria: Timing, Risk of Loss, and Behavioral Risk. Figure 7 gives an overview over the factors and the relative weight they take in the risk assessment.

Criteria	Quantitative Indicators	Weight
Timing Risk	1. Volatility of the Outperformance p.a.	25%
Risk of Loss	1. Highest moving loss over a period of 6 months in the last 5 years	20%
	2. Probability of one month with loss	10%
	3. Average Loss during a loss month	10%
	4. Negative Elasticity	20%
Behavioral Risk	1. Tracking Error	15%

Table 7: Risk Indicators for Feri Fund Rating

Timing risk frames the risk of generating high losses through wrong exit or enter strategies. On a quantitative basis this can be evaluated by variations of the returns, i.e. the annualized volatility of a funds outperformance compared to the index.

60% of a fund's overall risk assessment are however based on the risk of losing money. In the quantitative analysis four criteria will be considered in order to describe the risk of loss. Three of them are centered on the loss characteristics a fund shows: the highest 6-month moving loss in the last five years, the probability of loss during a month and the average loss during a loss month. The fourth criterion is the negative elasticity of a fund. This number shows how a fund behaves during market down-turns. It measures the percentage change in fund return when the market return declines by 1%.

The last criterion is the Behavioral Risk of a fund. This figure captures the risk arising from the active portfolio management. On a quantitative basis this can be captured by the tracking error, measuring how closely a fund follows its benchmark. The rater can therefore evaluate if additional risk arises from the portfolio management, since a high tracking error can have its origin in a high over- or underperformance of the fund.

5.3.2 FERI Fund Categories

Feri distinguishes 53 style categories in Germany. At least 20 funds have to be assigned to one category. Feri states that a reasonably large fund category is necessary to ensure that ratings are assigned based on true relative performance of the fund among its peers and not just by lack of competitors. The investor can also find a risk rank of the investment style categories. This rank is based on the risk classification of the European Securities and Markets Authority (ESMA), which ranks funds according to their risk profile from 1 to 7 (CESR, 2010). The different style categories can be found in the Appendix 5.

5.3.3 FERI Final Rating

In order to rate a fund, Feri compares its performance in each of the 12 indicators outlined in table 6 and 7 to its fund category. In order to arrive at the overall rating the quantitative indicators are then standardized and graded on a scale from 1 to 100. For each indicator the funds will get ranked according to their performance within the

peer group. In order to do so the median-ranked fund will be identified and receives 50 points. The extreme funds, i.e. the worst and the best fund within the peer group receive 1 and 100 points respectively. The remaining funds in the category score in relation to these funds. This is done for all 12 indicators. The final rating is then conducted by weighing all the points according to the percentage weight of the 12 indicators. Funds that generate at least 78 out of 100 points achieve an A-rating. Funds with a total score above 60 points achieve a B-rating. The grade C will be assigned to funds with a score above 41 points. A D-rating will be achieved by funds that score above 23 points and funds with a score below 23 will be E-rated. This rating distribution implies that a category might not show all rating grades from A to E. For example there might be categories without any A-rated funds at all, if no fund scores above 78 points. This rating distribution method is very different from Morningstar's and Lipper's, where the rating distribution within one investment category is fixed and always the same.

It is important to note that the rating is also of relative nature, meaning that a fund is evaluated in relation to the competitive products in the peer group. A comparison with funds outside the peer group is therefore not possible. The same is true for comparisons across different country borders. Feri's ratings are only specific to a certain country, as the competitors are different in the fund peer group of another country.

5.4 Overview Rating Methods

In order to give a comprehensive overview over the similarities and differences between the rating methods and to sum up the answer to the question *How do the raters process and present the available fund information in the ratings and what are the differences between the different fund rating agencies*, the following table 8 gives an overview:

	Morningstar	Lipper	FERI
Number of Ratings for a fund	One overall rating	Five different ratings	One overall rating
Symbol "Best Grade"	★★★★★	5 "Lipper Leader"	A
Rating Method	Risk-adjusted Return to peer group	Outperformance of peer group in 5 distinct rating merits	Outperformance of peer group based on Performance and Risk indicators
Number of rated funds in the German Universe	26,573[6]	16,888[7]	3,417[8]
Rating Calculation	Continuity of the risk-adjusted return over 3-year, 5-year, 10-year and overall horizon	Continuity of several criteria in five different ratings over 3-year, 5-year, 10-year and overall horizon	Outperformance based on Risk and Performance Indicators over a 5-year period.
Minimum Age	3 years	n.a.	5 years
Risk Considerations	Volatility, but emphasize on downside risk	H-exponent	Loss probabilities and characteristics, negative elasticity

---

[6] As of 07.05.2012 retrieved from
http://tools.morningstar.de/de/fundscreener/default.aspx?Site=de&LanguageId=de-DE
[7] As of 07.05.2012 retrieved from http://www.lipperleaders.com/index.aspx
[8] As of 31.05.2012 retrieved from: http://www.feri-fund-
rating.com/Default.aspx?Name=FundsRatingGermany&Content=TopFonds&Lang=en

Fixed Rating distribution within investment style category	Yes (according to normal distribution)	Yes (20%-steps)	No
Minimum Number of funds within one category	20	10	20
Dimension	Two-dimensional	Multi-dimensional	Multi-dimensional

Table 8: Overview Rating Methods

6 Components of Fund Ratings

Despite the increasing investment in mutual funds in the recent decade, individual investors still lack basic knowledge about the mutual fund market and are often unaware about specific features of their fund investment. Several studies have focused on this aspect and investigated the individual investor's financial expertise. The researchers Capon et al. (1996) documented in an U.S.-based study that only 60.7% of the mutual fund investors knew whether their funds were classified as load or no-load funds. Furthermore only one quarter of the investors was able to state the investment management style of their fund. A different study from Alexander, Jones and Nigro (1998) finds that only 43% of the investors had knowledge about the expenses their funds charged at the time of purchase. 20% of the investors also assumed that higher expenses were associated with higher returns.

This lack of knowledge gives a first indication why investors are interested in fund ratings, as they give a third-party evaluation of the funds based on different attributes of the investment. It is therefore important for this study to look into the different information components of fund ratings and their importance for helping the investors to evaluate and compare the quality of the different funds. In the next section we will therefore elaborate which factors should be incorporated and what drawbacks of these factors should be considered in a fund rating methodology in order to provide a valuable assessment of a fund's past performance. These factors include the Fund Categories, the Fund Costs, the Fund Risk, the Extreme Risk of a Fund and the Per-

formance Persistency, which should be considered when evaluating a fund's past performance. Moreover, the Comprehensiveness, Reliability and Transparency of a fund rating complete these factors, as they are important for the understanding of the individual investor.

6.1 Fund Categories

The mutual fund industry uses investment objectives and style classifications in order to categorize the differences between investment choices. Mutual funds are usually categorized by the type of securities the fund invests in and by the investment style the manager claims to follow (Brown and Goetzmann, 1997). With the tremendous increase in mutual fund numbers also the number of fund categories increased considerably. Therefore a distinction between industries and countries is also common today.

The style accounting has a major influence on the rating a fund achieves. The fund category determines the peer group among a fund competes and therefore determines the group of funds among a fund's relative performance gets measured. The importance of this style criterion gets visible when we look at studies that compare the impact of asset allocation and portfolio management on an investment's performance. Empirical studies (e.g. Brison, Singer and Beebower, 1991) found that the strategic asset allocation is the highest contributor to portfolio return, i.e. in which proportion the investor choses equities, government and corporate bonds, cash or real estate.[9] The specific security picking skills only contribute little to the portfolio return on average. If a fund is assigned to a very broad category or even to a 'wrong' category, the fund is not evaluated based on its management skills but only because a certain asset class was favored or not favored by the market throughout the evaluation period (Amenc and Le Sourd, 2007). The right category classification is therefore crucial in order to assess the past relative performance of a fund.

It is of special interest to look at the issues of fund categorization into more depth. Areas of concern are the broad and narrow categorization, the style evaluation by the manager and the case of indistinct fund categorization.

---

[9] Brison et al. (1991) come to the conclusion that over 90% of the variation in quarterly total return is attributable to the asset allocation.

6.1.1 Broad and Narrow Fund Categories

Fund raters are facing the trade-off between narrow and broad asset classes. On the one hand, narrow fund categories might build a too small peer group and might limit the diversification possibilities of fund managers among different asset classes, industries or countries (Amenc and Le Sourd, 2007). Broad asset classes circumvent this weakness and allow for more investment choice for the fund manager. Brown and Goetzmann (1997, p. 374) outline that broad definitions of mutual fund categories give "great latitude in the types of stocks to hold, the timing of purchases and sales, the level of fund diversification, the industry concentration of the portfolio, and a host of other factors that go into determining the returns to client investments". On the other hand broad fund categories might lead to the comparison of funds that are like apples and oranges and therefore do not account for the individual performance of a fund but its asset allocation as outlined above.

6.1.2 Fund Category Evaluation by the Manager

The process of assigning a fund to a specific fund category should be shaped by objectivity and not based on the fund category the investment manager states. This is based on the fact that in several cases there has been a misclassification of funds by the investment company. This might either happen through the assignment of a wrong investment category and benchmark index or through a misclassification in the fund name. DiBartolomea and Witkowski (1997) find in their research that 9% of all U.S. equity funds are misclassified and 31% have some misclassification in the categorization stated in the fund companies' prospectus. A study by Kim, Shukla, and Tomas (1995) comes to similar results stating that half of the 591 investigated funds state style classification in their prospectus that are different from their own objective assessment. The high number of misclassifications in both studies points to the conclusion that these misclassifications were not the effect of random noise. Instead diBartolomea and Witkowski (1997) argue that this result is based on „the competitive pressures in the mutual fund industry and compensation structures that reward relative performance" (p. 32). Brown and Goetzmann (1997) come to a similar conclusion and support the second reasoning from diBartolomea and Witkowski in their research work. In their paper from 1997 they find evidence that managers' misclassi-

fications were intentional in order to increase ex post performance measures and thereby the remuneration of these investment managers.

Objectivity in the categorization process should therefore be stressed for fund rating agencies since it lowers the moral hazard that managers face in determining their fund categories without objective criteria (Brown and Goetzmann, 1997).

6.1.3 Indistinct Fund Categorization

Furthermore, problems can arise when funds do not fit to the existing investment style categories or the investment style is not easily to detect. While distinguishing the funds based on their focus industry, country or asset class is less problematic, a distinction between value and growth shows more complications. Popular measures to distinguish value from growth strategies are the P/E, book-to-market and dividend yield ratio of the individual stocks the fund holds. These measures are however highly influenced by the market conditions. Risager (2009) for example shows in his book "Investing in Value Stocks" that during the dot.com bubble 1999 and the housing bubble from 2006 these ratios were on exorbitant levels, making it very hard to distinguish between value or growth characteristics of a stock and therefore of a fund. Additional problem is when innovative and new investment styles might not have enough comparable competitors yet and cannot easily be assigned to another group. In this case funds should not be assigned a rating.

6.1.4 Investment Style Persistence

Yet another important aspect for mutual fund raters to investigate is the style persistence of a mutual fund investment style. Evidence from the a very recent study of the U.S. mutual fund market shows that funds with lower volatility in their fund investment style on average outperform funds with a higher volatility in their fund investment style, even after adjusting for risk (Van Harlow, Brown and Zhang, 2011). Fund raters should take this into consideration when a fund changes the investment style and should put their current rating on hold and start a new rating after a certain time frame.

6.2 Fund Costs

The costs that are associated with an investment in mutual funds are of high im-

portance for the individual investor. Several academics showed that there is a strong relationship between mutual fund performance and fund expenses. Carhart (1997, p. 80) shows in his study that "expense ratios, portfolio turnover, and load fees are significantly and negatively related to performance". Morningstar (2010a) also conducted a study in this research field. They used expense ratios for U.S. funds in the U.S. asset classes domestic equity, international equity, balanced, taxable bond, and municipal bond for the period of 2005 through 2008 in order to predict future performance of the funds. Their findings show that for all time periods and data points low-cost funds outperform the high-cost funds.

However, as we noted above fund investors are sometimes not entirely aware of the fees they are facing when investing in a mutual fund or even associate higher costs with higher performance. For example a study over 30 years of U.S. mutual fund flows shows that investors are very sensitive to front-end loads and commission rather than to operating expenses (Barber, Odean and Zheng, 2003). As the investors might not be able to process the mere information about fund costs, mutual fund raters should therefore incorporate all costs in the evaluation process of past performance, i.e. the performance of a fund net of its costs.

6.3 Risk Assessment

According to the pioneer work of Jensen in 1968 there are two dimensions by which portfolio performance should be evaluated.

(1) The manager's ability to predict future security prices in order to increase the return on the portfolio.

(2) The manager's ability to reduce the inherent risk in this portfolio through efficient diversification.

The relationship between risk and return should therefore be in the center of the fund evaluation. Investors may increase their expected return by increasing the inherent risk levels, i.e. the probability of a loss (Modigliani and Modigliani, 1997). One commonly accepted reason for this is the risk aversion of the investors. As a compensation for uncertainty investors require a higher return, i.e. a risk premium over the risk-free rate.

Early work on the performance measurement theory already incorporated this knowledge and concludes that performance of a fund should be adjusted for the inherent risk in order to get comparable (Sharpe, 1975). However, there is still a lack of understanding of the nature and measurement of risk and return and its application in the mutual fund evaluation (Jensen, 1968). There has not yet been an overall agreement which measures capture the implied risk-return relationship best. Most widely used is the Sharpe ratio that measures the excess return over the risk-free rate divided by a risk measure, usually the classic past return standard deviation. Other measures like Jensen's alpha or the Treynor ratio incorporate most commonly the risk measure of the Capital Asset Pricing β in order to adjust the excess returns.

Besides the uncertainty on how to capture the risk of an asset, fund raters also have to consider that the risk of a fund may be changed over the period in order to participate on bullish movements and to reduce exposure during bear markets. Recent work on this topic found that funds change their risk significantly over time (Huang, Sialm and Zhang, 2008). They also find that "funds that increase risk perform worse than funds that keep stable risk levels over time" (p.1). Huang et al. (2008) argue that this is not due to portfolio managers' surge of market timing but opportunistic behavior, i.e. an agency problem. This principal-agent problem has been subject to several studies. Massa and Patgiri (2009) and Starks (1987) for example focus on the potential agency conflict that arises from asymmetric information between the investor and the fund manager. The investor does not know if the manager as the agent is truly pursuing the interest of the investor. It is therefore possible that the fund investor wants to increase their risk-adjusted security return, the mutual fund manager however rather desires to maximize his remuneration. The findings of the study from Massa and Patgiri (2009) are in line with this theory as they find that contractual incentives increase the risk taking of mutual fund managers. It is therefore in the interest of the investor that fund-rating agencies take the risk of a fund into consideration when constructing their rating.

6.3.1 Extreme Risk Assessment

The lesson from the financial crisis in 2007 taught us that the traditional risk management standards were not sufficient to handle all extreme risk. The bankruptcy of Lehman Brothers in 2008 showed the world that the probability of a total fallout is not

as small as anticipated. Investigation also found that the probabilities of these extreme events are even higher than anticipated by the normal distribution (Goldberg, Menchero, Hayes and Mitra, 2010). Indeed, the distribution of asset returns shows fat tails. Since researchers found that investors regard these small probabilities of very bad outcomes as a zero probability (Fischhoff, Slovic and Lichtenstein, 1977) these extreme events can not be neglected in a fund rating.

Extreme risk measures should therefore also be taken into consideration when evaluating the risk-adjusted performance of a mutual fund. Research already addressed these findings and introduces new measures to estimate this extreme risk. Examples are McNeil (1999) who investigates the extreme value theory as a method for modeling and measuring extreme risk. Rachev, Martin, Racheva and Stoyanov 2006) introduce the Stable Tail Adjusted Return Ratio in order to adjust the classic Sharpe ratio for the fat tails in the normal distribution

6.4 Performance Persistency

The investor fancies a fund rating that offers a continuity component in its methodology. A fund that showed consistent high returns should receive a better rating than a fund that has just performed well in the recent past. When evaluating the past performance of a fund, raters should therefore use a method to weigh the performance throughout the existence of the fund. Different time periods should therefore be taken into consideration when evaluating the performance of a fund, i.e. both long-term and short-term performance should be considered when evaluating.

6.5 Comprehensiveness, Reliability and Transparency

Several research studies elaborated investor's lack in knowledge of the mutual fund market in general and about their own investments in specific. Investors therefore have a need for a reliable information source. As we showed above ratings contain a lot of valuable information content for the investor. Raters urge to deliver this information in a comprehensive and understandable way. However, they also assure the investor that this information is correct and not influenced by the mutual fund firms. In case of credit ratings there is always the discussion that the dependence between rating agencies and rated firms is too high, since firms pay a non-negligible fee in order to get a security rated (SEC, 2003). In contrast to credit ratings the independ-

ence of the raters is more reliable. Weather a mutual fund gets rated does only depend on the raters and is completely free for the fund provider. If there is reliability in information accuracy by the third-party, the rating can be a valuable asset for the investor.

Transparency is however also an important point for the mutual fund rating's value. Investors that are interested in the construction of the ratings should be able to gain this information in order to evaluate the individual information value a specific rating offers them. One investor for example might be interested in the extreme risk assessment a rating offers, others consider this as an obsolete feature of a rating and might prefer one that does not take it into consideration. Therefore it is important for the rater to disclose this information and make it accessible to the investor.

7 Fund Rating Analysis

In the following part the mutual fund methods of Morningstar, Lipper and Feri will be assessed based on the above outlined factors in order to answer the question: *Where do the offered ratings show weaknesses and strengths in helping the investor to assess funds' past quality (performance)*. Furthermore, problems with the Age Bias and the Percentile Distribution of Morningstar and Lipper will be included, as they are also decisive for the quality of a mutual fund rating.

7.1 Fund Categories

As outlined above fund categories can have a significant influence on the relative performance of a fund and therefore the rating of a fund. The three raters have different approaches to deal with this problem, which will be assessed in the following. Before 2002, Morningstar assessed a fund compared to very broad asset classes. Funds were classified into the four categories, domestic equity, international equity, taxable bond, and municipal bond. However, Morningstar introduced a new categorization for the U.S. in 2002 and for Europe in 2006. The newly introduced fund categories are narrower, mainly because the former broad asset classes did not take the investment style of a fund into account. This step was motivated by several studies that tried to explain differences in return with different investment styles in terms of market capitalization and value/growth properties of the stocks, as outlined above. The narrow fund categories have the advantage that the ratings assess the added

value of an investment manager and not only the temporary attractiveness of an investment style. Morningstar argues that now the funds within one category share the same risk properties and can therefore be considered as a substitute to each other. In order to cope with the increasing number of different funds Morningstar continued to increase the number of categories for Europe dramatically in the recent years. For the German fund universe there are now more than 270 categories for 26,918 funds available. (Morningstar, 2012b)

Morningstar decides about the fund category based on the underlying securities the fund invested in over the last three years. Fund classifications might therefore be different from the stated category of the investment company. However, the fact that financial products get more and more complex over time and that Morningstar is only able to assess the holdings which are proclaimed by the investment manager increases the possibility that risk profiles of funds might not completely be alike.

Lipper is also pursuing a strategy of rather narrow fund categories, as they provide 289 categories for 17,062 funds. However, Lipper does not take investment style into account when assigning a fund to a category. Lipper only uses the asset type, market capitalization, country, region and industry focus. Different from Morningstar, the main input factor for the fund category decision is the classification provided by the investment company. This standard can be considered critical in two ways. First, Lipper does not take the performance differences of investment styles into account. This is problematic, because, for example, there have always been periods in which value stocks outperformed growth stocks and vice versa (Risager, 2009, Chan and Lakonishok, 2004). In Lipper's fund classification value and growth funds might therefore get compared to each other, which will not lead to a valuable comparison of past performance. Style Accounting should be considered when assigning a relative rating to a fund. Second, the main input is the investment objective stated by the investment company itself. In the light of several studies that give evidence for high numbers of fund misclassifications by investment companies this does not seem appropriate. However, Lipper is aware of this problem and monitors the outlaying funds within a fund classification in order to assess the appropriateness of a classification (Lipper, 2008).

Feri on the other hand rated 3.185 funds in 54 categories in 2010. Compared to

Morningstar with currently 275 categories and 26,918 funds and Lipper with 289 categories and 17,062 funds, Feri rates much less funds in much less categories. One reason for this is, that Feri only rates one share class of a fund compared to Morningstar and Lipper, who rate several share classes and therefore have a higher average number of funds in their categories. The limited fund universe of Feri however indicates that the fund categories have to be broader than in Morningstar's and Lipper's case.

In order to assign the funds to different categories, Feri uses only the country- and industry- focus and the TOP 10 holdings stated by the investment company (BVI, 2004). Other than Lipper, Feri is therefore not using the category stated by the investor as a main input factor but also uses the holdings of the fund to verify this information. However, because of the limited fund universe, Feri distinguishes not between market capitalization and investment style. In the same way as for Lipper this might be critical as there have been times where certain investment styles including market capitalization have outperformed others.

Morningstar therefore shows a clear advantage in the fund classification process over Feri and Lipper, as it offers narrow fund categories with value and growth characteristics and market capitalization, but also uses the holdings of a fund to verify the fund category.

7.2 Fund Costs

As outlined above fund costs are a very important determinant for the performance of a fund. Morningstar and Lipper use different methods in order to evaluate a fund's expense structure.

Morningstar adjusts the return of a fund for expenses in the second step of its MRAR calculation. This leads to the load-adjusted return:

$$LR_t = (1 + TR_t)(1 - F)(1 - R_t) - D_t(1 - F)\frac{\min(P_0, P_1)}{P_0} - 1$$

First, Morningstar adjusts for the operating expenses implicitly, as mutual fund returns after operating expenses are used for any calculation. Second, Morningstar takes the front load (F), the redemption fee (R_t) and the deferred load (D_t) into consideration. The deferred load is applied to the smaller of the NAV at the beginning or at the end of the period, which is in line with the industry standard (Morningstar,

2009). If the deferred load is straddled, Morningstar uses the higher one. Morningstar does not consider the historical fee structure, nor does it adjust for the rating if the fees only change on a temporary basis. The load fee is capped at 5% and at 3% for equity and fixed-income funds respectively. In theory this is in line with the stated load of the distributors. However, after the financial crisis "the actual amounts charged in practice are often much lower or waived entirely" (EFAMA, 2011, p. 12). Morningstar might therefore adjust for loads that the investor is not facing in reality. Since the step of load adjustment is not reproducible for the investor, the entire Morningstar rating can be considered as biased, if the investor is facing other costs than anticipated by Morningstar.

Lipper employs a separate rating for the expenses of a fund. It groups funds within one fund classification according to their different load structures. Funds with similar load structures are then ranked according to their expense ratios and are then rated based on the 20% distribution outlined above. The Lipper Rating of Expenses does not take any returns of the fund and therefore no operating costs into consideration but advises the investor to use the Expense Rating in conjunction with another performance rating, like the Rating for Consistent Performance or the Rating for Total Return, where Lipper adjusted for operating expenses. A pure use of the Lipper Rating for Expenses might therefore be misleading when assessing the expenses of a fund, since only the loads of a fund are considered and operating expenses like management fee are omitted. Furthermore, investors should be aware that fund expenses vary a lot across all available fund categories. According to a study of the ICI (2011) the fund expenses depend on the asset class, fund investment objectives and other factors. For example equity funds tend to have higher expenses than bond and money market funds. Therefore Lipper's expense adjustment shows some advantages, as it does not assess the expenses on an individual basis, but compared to a similar peer-group. Since expenses vary among fund categories and investment styles it is a helpful procedure to compare the fund loads directly among peers.

In contrast to Morningstar and Lipper, Feri does not take any loads into consideration and only adjusts for the operating expenses in the return data of the fund. Therefore they neglect the influence of loads on the fund performance entirely when conducting

their ratings, which can take up a big part of the funds performance with loads up to 5% or 3% of the investment amount.

7.3 Fund Risk

The three raters use very different tools to adjust for the risk of a fund. In the following the methods of the three raters will be evaluated in more depth.

7.3.1 Morningstar

Morningstar uses a risk measure that is based on investor preferences. It is supposed to reward return and penalize risk at all times (Morningstar, 2009) and motivated by the Expected Utility Theory, which is a model to frame decision-making under uncertainty. The expected utility of a portfolio is defined as E[u(W)], where W is the ending value of the portfolio and u(•) is the utility function of the investor. The expected utility theory has the advantage that is does not rely on any assumptions about the distribution of returns, i.e. no lognormal or normal distribution is required. This is especially appealing since there have been numerous findings that asset returns tend to be leptokurtic (Mandelbrot, 1963, Fama, 1965 and Chunhachinda, Dandapani, Hamid and Prakash, 1997).

Before 2002, Morningstar used to assess the risk by looking at the degree to which mutual fund's returns underperformed those of a 90-day T-Bill. However, this method is only taking the downside risk of a fund into consideration. After the dot.com bubble in the 90's and 5-star rated Technology funds that collapsed, Morningstar saw the need to also measure volatility of upside returns. The old method failed to identify mutual funds that created enormous returns by risky behavior. The new risk methodology ensures that risks that have not yet surfaced because returns were only volatile in the upside. Therefore, funds with a high consistency get now awarded the best rating and for a high-risk fund it got harder to earn a five-star rating. Re-adjusting the ratings for the beginning of 2000 shows that with the MRAR method top-rated high-risk funds would have experienced a significant lower rating (Kräussl, 2007).

However, also the new method is viewed critical. There is discussion if the risk aversion factor of γ = 2, which is only based on the experience of Morningstar, is an appropriate approximation for investors' risk tolerance. It is questionable if the average risk aversion of all investors can be pictured at all (Weingärtner, 2009). Moreover,

there is a working paper by Lisi (2010), which investigates the risk accounting in the Morningstar rating. The researcher compared the actual ratings where $\gamma = 2$ with artificial Morningstar ratings where $\gamma = 0$, i.e. a risk-indifferent investor. The analysis shows that there is almost no difference between the two ratings, indicating that Morningstar does not appropriately account for the inherent risk of a fund. However, the study was not published and the academic methods were therefore not tested for their accuracy.

7.3.2 Lipper

Lipper has two ratings that measure the risk of a fund, namely the Rating for Consistent Return and the Rating for Persistence. The Rating for Consistent Return uses the theory of the H exponent to adjust for the fund risk. In this context the H exponent measures the smoothness of fund returns. The calculation is very complex but has some inherent drawbacks. There are several ways to construct the H exponent. The R/S analysis was the most common construction method of the H exponent (Qian and Rasheed, 2004). However, the analysis shows some shortcomings:

(1) It is only robust for large sample sizes.

(2) It shows a short-range dependence in the time series.

In small sample sizes the variance of these sample sizes are very large. As Lipper does not state a minimum fund age in its methodology, this can become critical, when rated funds have too short track record.

However, Lo provided a modified R/S to estimate the H exponent. Teverovsky, Taqqu and Willinger (1999) examined this modified version with synthetically generated time series. They find that the method does not provide a reliable test for the long-range dependence, as it tends to accept the null hypothesis of no long-range dependence. They claim that at least two different computations of the H exponent are necessary. The use of the H-Exponent might therefore show some inherent drawbacks and might not be able to depict the real risk behavior of a fund.

Lipper's Rating for Persistency shows another feature for the risk assessment. As it measures the months in which a fund incurred losses it is a measure for the extreme risk of a fund and can be used by investors that are most risk averse in the region of losses.

$$LP = \sum_{t=1}^{n} Min\,(0, R_t)$$

The highest number a fund can achieve is 0. The number will go down either if the number of loss months or the incurred loss in a loss month increases. The rating is therefore a measure of persistency of returns and gives the investor an understanding of the loss characteristics of a fund. However, the track record of a fund should cover an appropriate time frame where both bullish and bearish phases occurred. If the fund is too young and only covers one market phase the Rating for Persistency can be misleading.

Furthermore, it only assesses the downside persistency. As outlined by Morningstar, this would lead to assigning high ratings to funds that incorporate a high level of risk, but showed high return because their investment style or asset class is currently 'hot' and can therefore clearly be seen as a downside for Lipper's risk assessment.

7.3.3 Feri

Feri uses multidimensional risk indicators that account for 30% of the fund rating. The quantitative indicators include: volatility of the outperformance per annum, the highest moving loss, the probability of one month with loss, the average loss during a loss month, negative elasticity and the tracking error. This method is clearly incorporating several different aspects of the fund risk and can therefore be considered as multidimensional. Furthermore, Feri employs two risk measures that show important features, namely the Tracking Error (TR) and the past volatility. TR treats the variability of positive returns the same way as the variability of losses and therefore punishes positive outperformance of the fund (Clark, 2003). Moreover, volatility also punishes positive outcomes of the performance indicators. As outlined above this is important in order to identify mutual funds that created enormous returns by risky behavior. However, historical measures of the volatility of the benchmark outperformance should be treated as critical, as they might be influenced by some extreme events that do not show the intrinsic risk of a fund.

Another feature of the Feri risk measures shows a considerable advantage. 60% of the risk assessment are based on the loss behavior of a fund. Therefore, extreme risk assessment of a fund is accounted for. As Feri also requires a minimum track

record of five years, the accuracy of these extreme risk measures is more likely than in Lipper's case.

7.4 Comprehensiveness, Reliability and Transparency

While Morningstar and Feri provide one overall rating for each fund, Lipper assigns up to five different ratings to one fund. On the one hand the fund ratings might therefore lack the characteristic of comprehensiveness, since the investor first has to read the description of each rating. On the other hand the policy of five different ratings gives the investor more insight into the relative performance of different features of the fund. The investor is now able to see in which merit the fund has its weaknesses and its strengths, e.g. in risk assessment or the expenses charged. It therefore offers further information to the investor.

Another aspect to consider is the reliability of a rating. An investor demands the independence of mutual fund raters. Lipper and Morningstar do not engage in any financial management activities, but provide different products of fund and financial information to individual and professional investors. However, Feri EuroRating Services AG and its holding Feri AG are both entirely owned by MLP Finanzdienstleistungen AG (MLP, 2011), a German financial advisor for professional and individual investors. The company administrates more than EUR 20.2 billion in investments solely in Germany (MLP, 2012). Therefore, a possible manipulation of the ratings could be advantageous for MLP. Although the organizational separation might prevent most interference, this is certainly a drawback of the organizational constitution of Feri.

The transparency of a rating is also an important point for the quality of a rating. All methodologies are available on the Internet and therefore accessible for the individual investor. However, in case of Feri the methodology is in German language and therefore not available for non-German speakers.

7.5 Percentile Distribution
While Feri does not assign the rating based on a fixed distribution, Morningstar and Lipper both use a predefined percentile distribution for each fund category.

	Morningstar	Lipper
Percentage of Funds within the three best rating grades	67.5%	60%
Position of the third best rating grade	Between 32.5% and 67.5%	Between 40% and 60%
Position of the last category	Bottom 10%	Bottom 20%

Table 9: Percentile Distribution of Morningstar and Lipper

Compared with Lipper, Morningstar assigns the best ratings to a higher fraction of the funds. The investor might therefore think that a higher number of funds performed well during the past, than in Lipper's case. Feri does not employ a fixed percentile distribution but reports that around 8% of all funds receive the highest rating of A. As the raters therefore employ different standards for the percentile distribution, the same rating grade, e.g. five-stars for Morningstar, 5 for Lipper and A for Feri, does not indicate the same relative standing of the fund (Amenc and Le Sourd, 2007). This becomes especially clear when looking at the position of the third best rating grade of Morningstar and Lipper as outlined table 9.

7.6 Age Bias

As outlined above, Morningstar defines the overall rating as a weighted average of the three-, five- and ten-year ratings. There has been research investigating if this weighing method is favoring the appearance of an age bias. Blume (1998) was the first one to investigate the age bias of the Morningstar rating. He found that seasoned funds are less likely to earn a higher rating. Morey (2000) on the other hand found that younger funds are less likely to earn a higher rating. Even though the results of these papers are quite different, both attribute the age bias to the weighting system of Morningstar. Even after the significant methodology change in 2002 Adkisson and Fraser (2004) still find an age bias. The scholars find that there is an inverse relationship between fund age and fund rating. When assessing the weighing system of

Morningstar it becomes clear that the three-year rating contributes a major part of every fund age. For the calculation of a seasoned fund (older than 10 years) the most recent 36 months are involved in the calculation of the three-year, the five-year and the ten-year rating. Based on the outlined weights the actual data impact of the first 36 months is therefore 53%. For a middle aged fund (less than 10 years, more than 5) the actual weight even increases to a total of 76%.

However, Morningstar is aware of the possible age bias and it is intentional that the last 3 years make up for such a great part in all ratings. Morningstar states that the last three years take up such a major part, because major changes in the portfolio management and the environment might have occurred (Morningstar, 2009). Because Morningstar still wanted to take them into account in order to acknowledge long time performance they do not weigh the 10-year, 5-year and 3-year rating equally. The method is however still questionable as management changes can also be communicated in a different way, which will be outlined in part 7.7.

Lipper rates the age ratings equally over the overall period. This could as well lead to an age bias in the ratings. However, Lipper also indicates the different age ratings on their database, so that the investor is aware of the actual performance of the fund over the years. Morningstar however does not provide the investor with this information and only states the overall fund rating.

Feri on the other hand does not apply any weighing of different ratings over the years. The quantitative indicators are solely received from the last five years of past performance, no matter how old the fund actually is. Since the age bias is mainly attributable to the weighing there is most likely no age bias in the rating.

7.7 Management and Style Changes

As outlined above it is very important for the investor that a rating gives an assessment of a fund's capabilities. If the management or the investment focus of a fund changes, the investor should be informed because this might also affect the performance of a fund. However, Morningstar and Lipper do not handle a fund differently when the fund management changes, as in the case of Feri. If the portfolio management changes, the rating is marked with „ur" (under review), indicating that the fund has undergone some recent changes in either the fund management or the portfolio composition. The fund is on the watch list of the rater for 18 months until the invest-

ment style is confirmed or the rating will be reviewed. Morningstar only suspends the rating of a fund if it has changed broad asset classes, e.g. international equities, taxable bonds or it was holding 100% cash over a time frame of more than one year (Morningstar, 2009). After three years without category changes the fund is eligible for a new rating.

Every rater faces a trade-off between making ratings too volatile and the steady incorporation of new information of the fund. If the fund quality is experiencing changes the fund rater should incorporate these improvements or drawbacks, as the investor is losing money with every time lag. However, the ratings should not be too volatile. It is therefore in the sense of the rater to adjust a rating only if important information with actual impact is coming up.

7.8 Evaluation Summary of Methodologies

The above analysis sought to answer the question: *Where do the offered ratings show weaknesses and strengths in helping the investor to assess funds' quality?*

Fund raters show very different benefits and limitations. While Morningstar shows the best solution for the Categorization of funds, it shows weaknesses in its applied Percentile Distribution of the rating grades and the possible age bias for an adequate assessment of past performance. In the same way Lipper shows some weaknesses in the percentile distribution and might also be subject to an age bias. However, Lipper shows clear strengths in the assessment of fund costs and the Comprehensiveness, Reliability, and Transparency of their rating. Contrary to that Feri has its weaknesses in the fund cost evaluation and the Comprehensiveness, Reliability, and Transparency of their rating. Clear strengths are the handling of age or percentile distribution bias and in case of management or investment style changes. These findings are summed up in the following table 10.

	Morningstar	**Lipper**	**Feri**
Categories	- narrow categories - categorization is subject to biases because of limited information on holdings	- narrow categories - no adjustment for the style or market capitalization of a fund - categorization is subject to biases, information of asset manager main input factor of categorization	- broad categories - no adjustment for the style or market capitalization of a fund's holdings - categorization is subject to biases because of limited information on holdings
Assessment	Good	Fair	Fair
Fund Costs	- adjustment for loads - operation and management fees accounted in return data	- own rating for expenses - expenses get compared to funds in same category - big differences in expense standard between categories	- operating and management fees accounted for in return data - no adjustment for loads
Assessment	Good	Good	Poor
Risk	- accounting for upside risk - risk aversion factor of 2 is questionable	- two different risk measures - extreme risk accounting - no accounting for upside risk	- multiple risk dimensions - extreme risk accounting - accounting for upside risk
Assessment	Good	Fair	Fair

Comprehensiveness, Reliability, and Transparency	- information material available online and in several languages	- information material available online and in several languages - several ratings offer additional information	- information material is only available in German - owned by a financial advisory firm
Assessment	**Good**	**Good**	**Poor**
Percentile Distribution	- fixed percentile distribution misleads investors about absolute fund performance	- fixed percentile distribution misleads investors about absolute fund performance	- no fixed percentile distribution
Assessment	**Fair**	**Fair**	**Good**
Age Bias	- age bias because of weighing method	- possible age bias because of weighing method	- no weighing method
Assessment	**Fair**	**Fair**	**Good**
Management / Category Change	- only broad category changes get suspended and marked		- funds with management changes get marked and on watch list
Assessment	**Good**	**Poor**	**Good**

Table 10: Evaluation of the Methodologies of Morningstar, Lipper and Feri

When answering*: What is the information for the investor received by the fund rating* several aspects have to be considered.

Each of the providers gives the investor an understanding of the relative quality of the fund, measured by its past performance. This service can be used for fund screening, as it gives a comprehensive overview of the features of a fund. However, as most ratings are only comprised in one "number", it is hard for the investor to look into the details of the fund features. For example the expenses and relative risk standing of a fund is not accessible for the individual investor. The use for fund screening that is stated by the rating providers is therefore only based on the relative past performance of a fund measured by their individual methodology. A notable exception forms Lipper, which gives five different ratings with different evaluation metrics to a fund. The ratings are still only a relative comparison but give an indication for different aspects as the fund expenses, fund risk and the extreme risk of a fund.

When assessing the individual ratings of Morningstar, Lipper and Feri, one fact became clear: the outlined construction of the ratings clearly suggests a link between rating and performance to the individual investor. The individual features of the fund are not presented to the investor in detail, rather one single number gives an assessment of past performance and the detailed features of the fund are hidden in the complex rating construction. Given the predictive characteristics of credit ratings, the investor might therefore assume a similar link between fund rating and fund performance.

Furthermore, the investor has to bear in mind, that all three rating providers show weaknesses in the objectivity of the fund categorization, which leaves room for manipulation by the fund management. Fund Managers still have the possibility to influence the rating, for example by changing their fund style, in which case their track record gets compared to a different set of funds in another category. In this case the quality of the fund management gets compared to funds that were facing a different market environment. The ratings can therefore not be considered as a accurate assessment of a fund's features. This is intensified by the fact, that the investor has no overview over the features of a fund and remains uninformed why a fund received a certain rating.

8 Regression Analysis of the Performance Predictability

In order to give a full understanding of the value of mutual fund ratings it is now important to solve the sub-problem of predictability. As outlined above several studies in the U.S. showed some investors understand the mutual fund ratings as a predictor for future performance. Fund Raters are well aware of the fact that their ratings are often consulted in order to make an investment decision. Some raters therefore clearly indicate on their websites and investment information material that their ratings are not a guarantee for a fund's future performance and rather an achievement score of a fund's past performance (Morningstar, 2009). There are two reasons why the raters might state this. On the one hand fund raters want to assure themselves against liabilities on a legal basis. On the other hand they know that future performance does not resemble the past, which is in line with the efficient market theory (Fama, 1970). An efficient market is defined as a market that incorporates all available information into the prices of the securities (Fama, 1970). Therefore no one should generate excess return by the use of public information, i.e. past returns.

However, when looking at the return predictability of mutual fund performance, we find mixed results. The performance predictability of mutual funds was subject to several studies, e.g. Hendricks, Patel und Zeckhauser (1993), Elton, Gruber and Blake (1995), Carhart (1997). These researchers examined the persistence in mutual fund performance. The results on this issue are not clear and make the general question of return predictability controversial.

Hendricks et al. (1993) were among the first to discover a short-term predictability for the 1975-1988-period among U.S. mutual funds. The simple strategy of quarterly selection of the top performers from the last four quarters significantly outperforms the average mutual fund. This phenomenon of return predictability was called "hot hands". They also find evidence for the counterpart phenomenon "cold hands" meaning that worst performing funds continue to underperform in the near future.

Elton et al. (1995) expand this study and find a predictability of future return from past performance for the short run as well as for the long run. They conclude that past performance contains information relevant for the future. This not only reconfirms the "hot-hands' phenomena but also shows a longer persistence of performance predictability in the future than anticipated by Hendricks et al. (1993).

Carhart (1997) however argues that this "hot hands" phenomenon is mostly driven by the one-year momentum effect, differences in the expense ratios and transaction costs. Therefore they conclude that there is no predictability in past performance and that their results are 'consistent with the market efficiency" (p. 80). However, their regression finds no explanation for the "cold hands", i.e. the only persistence they cannot explain is high underperformance of the funds with the lowest return.

A clear-cut answer to the question of performance persistence is therefore not possible. We therefore take a closer look at the return predictability of mutual fund ratings.

8.1 Data

In order to assess the predictive power of the raters, a regression analysis will be applied to assess the relationship between fund performance und past ratings in Germany. This analysis is based upon the methodology of the most robust and acknowledged study by Blake and Morey (2000).

8.1.1 Performance Data and Historical Ratings

The data for this study was retrieved from Morningstar and Feri. Data of mutual fund performance from Morningstar is generally available for researchers on a proprietary basis. Morningstar provided the arithmetic monthly returns of the Net Asset Value (NAVs) of all mutual funds that were available for sale in Germany from January 2005 to December 2011. The NAVs were calculated based on the total fund value adjusted for all expenses, divided by the number of outstanding fund shares. Furthermore, it was assumed that all dividends were reinvested. The data set of monthly historical overall fund ratings for Feri and Morningstar from December 2004 to December 2010 was provided from the raters Morningstar and Feri.

In order to make results comparable to previous studies the data was skimmed to equity mutual funds that have a German investment focus and were for sale in Germany at the point of time the rating was conducted. Only one share class per fund was assessed in the analysis in order to avoid double counting of data points.

The tested time horizon lies between 2004 and 2012, as it covers bullish as well as bearish market trends. It is therefore possible to test how the ratings perform under the different stages of the economic cycle. For Morningstar this procedure covers the categories "Germany Large Cap" containing around 70 funds and "Germany

Small/Mid Cap" containing around 30 funds over the years from 2005-2011, which is outlined in detail in table 11. The average Morningstar Rating over this period was 2.980 with a standard deviation around 1.02.

Year	Total Number of Funds	Number of Large-Cap Fund	Number of Small/Mid-Cap Funds	Average Fund Rating	Standard Deviation of Rating
2005	94	71	23	2.043	0.97181
2006	96	71	25	3.604	0.88828
2007	102	73	29	2.892	1.11610
2008	98	70	28	3.092	1.06574
2009	97	70	27	3.082	1.05742
2010	91	62	29	3.066	1.01983
2011	87	56	31	3.080	0.985196
Average	**95**	**67.4**	**27.6**	**2.980**	**1.01491**

Table 11: Morningstar Data Overview

For Feri this covers the Category "Equities Germany" with around 60 funds over the years from 2005-2011, which is outlined in table 12. The average Feri fund rating over this period was 3.023 with a standard deviation around 1.12.

Year	Total Number of Funds	Average Fund Rating	Standard Deviation of Rating
2005	60	3.150	1.23268
2006	64	3.000	1.16837
2007	61	3.098	1.12106
2008	63	3.000	1.12181
2009	61	3.082	1.14448
2010	54	2.963	1.04544
2011	52	2.865	1.01032
Average	**59.143**	**3.023**	**1.12060**

Table 12: Feri Data Overview

Feri also provides a mixture of qualitative and quantitative rating for funds younger than five years. Since the focus of this study is to assess the performance predictability of quantitative ratings, these fund rating were deleted from the list.

8.1.2 Adjustment for Loads

Blake and Morey (2000) used two different data sets for the performance of the mutual funds. One was adjusted for front-end and deferred loads, the other one was not adjusted for any loads. The reasoning behind the adjustment for loads is that Morningstar also adjusts for these expenses when conducting their Star rating. Feri however does not adjust for any loads and an adjustment for loads would therefore bias the results for Feri. Furthermore, there was no significant difference in the results of the analysis by Blake and Morey between the two data sets. The performance data for the predictability regression will therefore not be adjusted for any loads.

8.1.3 Performance Indicators

The magnitude of literature on mutual fund performance measurements indicates the importance of different performance evaluation methods in this study. There have been several approaches to measure the performance of a fund, like geometric mean, Treynor's ratio, Sharpe ratio, Jensen's alpha and Fama and French's three-factor model (Modigliani and Modigliani, 1997). Potential bias in the performance measurement cannot be avoided, since the question about the best performance measurement of mutual funds remains unanswered. However, the focus of the analysis will be whether there is any relationship between a fund rating and the subsequent performance of the fund, measured with the most common performance measures (Gerrans, 2006). The analyzed performance measures of this study will therefore include the most commonly used ones, which were also in the focus of previous studies. Past research has always applied different performance measures in order to assess the relationship between past ratings and performance. These measures include the geometric mean monthly return and the Sharpe ratio.

Geometric Mean Return

The geometric mean is a measure that is used to best describe the change in wealth over multiple periods (Jones, 2009). When assessing the performance of a fund,

scholars therefore prefer it to its arithmetic counterpart, which is better for single-period evaluation. The geometric mean is defined as:

$$r_{geometric} = (\prod_{i+1}^{n}(1 + r_{arithmetic,i}))^{\frac{1}{n}} - 1$$

where $r_{arithmetic,i}$ is the arithmetic return of the i-th fund and n is the total number of periods.

The second performance measure is the Sharpe ratio. The Sharpe ratio is a very popular measure to evaluate mutual fund performance (Gerrans, 2004). It is a risk-reward ratio, as it measures the excess return of a fund over the risk-free rate per unit of deviation (Jones, 2009). It is generally defined as follows:

$$S_i = \frac{R_i - R_f}{\sigma_i}$$

where R_i is the mean return of a fund i, R_f is the risk-free rate and σ_i is the standard deviation of fund i. However, in a paper from 1994 Sharpe revised his original ratio due to the fact that the risk-free rate and its approximation is not constant over the evaluation period. The revised historic Sharpe ratio is therefore defined as follows:

$$S_i = \frac{E(R_i - R_f)}{\sqrt{var(R_i - R_f)}}$$

In this study R_i will be the geometric mean monthly performance a fund. Further-more, σ_i will be the monthly standard deviation of a fund. The risk-free rate will be approximated by the 1-month EURIBOR rate of the respective month. The use of interbank rates as an approximation has been common practice in the European market place (Vaihekoski, 2009). As the 1-month EURIBOR arithmetic rates are pub-lished on a per annum basis, the rates will be transferred to a monthly basis in order to compare it to the monthly return and monthly standard deviation of a fund. Hereby it is important to note that the interest rate is quoted on an actual/360 day count ba-sis, which leads to the following calculation of monthly risk-free rate:

$$R_{f, \; monthly} = R_{f, \; yearly} * \frac{30}{360}$$

8.1.4 Dealing with Survivorship Bias

In the light of possible fund mergers and liquidations, survivorship bias has to be considered in the fund sample. Survivorship bias in the performance assessment is defined as the tendency to exclude liquidated and merged funds from the analysis and thereby introducing a higher skewness of the results (Brown, Goetzmann, Ibbotson and Ross, 1992). Brown et al. (1992) show that a data set with survivorship bias properties is more likely to show predictability. Studies of mutual fund ratings that do not take the survivorship bias into account might therefore lead to a false conclusion regarding the predictability of these ratings. German Equity funds that were liquidated, merged or changed names in the time period of 2004-2011 are therefore subject to survivorship bias in this study.

Table 13 gives an overview over the liquidated funds during the examination period. The average rating of liquidated and merged funds is below the average rating for all funds. Omitting these funds from the data would therefore introduce an upward bias of the performance data.

Rater	Number of Closed/Merged Funds	Average Fund Rating of Closed/Merged Funds	Median Fund Rating of Closed/Merged Funds
Morningstar	37	2.32432	2
Feri	21	2.28671	2

Table 13: Closed/Merged Fund Data

Following the method of Blake and Morey (2000), liquidated or merged funds should therefore not be omitted from the data. Blake and Morey (2000) introduce different methods to deal with this survivorship bias. The procedure for the adjustment for merged funds is straightforward, as the return data of the new, combined fund was used in the analysis after the fund merger. However in cases of fund liquidation three possible procedures are described in Blake and Morey (2000).

(1) The arithmetic mean of fund's past performance is used for the missing data points.

(2) A reinvestment in a similar fund with a similar rating in the same category is assumed.

(3) A balanced reinvestment in the remaining funds within the same fund category was assumed.

As any assumptions about the reinvestment decision of the investor of a liquidated fund to a particular rating grade or category might introduce a bias, a balanced reinvestment in the remaining funds of the original category is considered as the best solution. In order to adjust for survivorship bias in this study option (3) was therefore used. The missing return data of liquidated funds was filled with the category average. This allows the use of out-of-sample performance of the funds to avoid survivorship bias.

8.2 Methodology

In line with Blake and Morey (2000) and Morey and Gottesman (2006) fund ratings are taken from the end of the previous year. Performance measures over the subsequent one-year, three-year and five-year periods are then calculated. These performance measures from different observation windows are then regressed based on the following formula:

$$S_i = C + \gamma_4 D_{i,4} + \gamma_3 D_{i,3} + \gamma_2 D_{i,2} + \gamma_1 D_{i,1} + \mu_i$$

Where S_i is the performance of the fund over the evaluation period. D_1 through D_4 are the dummy variables representing the rating grades from one to four stars for Morningstar and from E to B for Feri. C is the performance of the reference group, i.e. the performance of a five-star fund and the performance of an A-rated fund respectively. γ_1 through γ_4 are the coefficients of the dummy variables. μ_i is the error term of the regression analysis.

In order to show predictability in the fund rating, the coefficients should show several attributes in the regression.

(1) Since the reference group denotes the highest rating grade, the coefficients of the dummy variables for the lower rating grades should all be negative.

(2) In the case of distinct rating grades the coefficients should show the following pattern: $\gamma_4 > \gamma_3 > \gamma_2 > \gamma_1$

The regression will be conducted at different points in time in order to test the predictability in different economic settings, i.e. bull and bear markets. The significance of the results will be tested with the use of t-statistics.

We will regress different time periods of one, three and five years in order to test for the short- and long-term predictability of the ratings. As Blake and Morey (2000) outline these time periods are also consistent with the past performance periods provided to individual investors when making the investment decision. No other independent variables are included, as the R^2 and the F-test of the regression analysis would be biased.

8.3 Results

The results of the regression analysis will be outlined in the next section. In order to detect heteroscedasticity in the data, the White test (White, 1980) was used. When the test indicated heteroscedasticity at the 10%, 5% or 1% level heteroscedasticity-consistent standard errors and t-statistics were computed in order to test the significance. Since the regression analysis was in the form of a least square dummy variable regression analysis, the t-statistics do not indicate the Null-Hypothesis $H_0: \gamma_i = 0$ but $H_0: \gamma_i - \gamma_{dropped} = 0$ (Park , 2002). Therefore, a significant t-test statistic tells that the coefficient is significantly different from the reference group, i.e. the five-star rating.

8.3.1 Morningstar

The results for Morningstar are outlined in table 14 and 15. The coefficients for the dummy variable coefficients and the constant are reported with their corresponding t-statistics. Furthermore, the R^2 and the F-stat of the overall model are reported in the last two columns.

Geometric Monthly Mean

In general, results only predict poor predictive powers of the Morningstar ratings. The F-values of the different evaluation windows indicate whether the independent variables are able to reliable predict the dependent variable. In 5 out of 7 models in the 1-year sample indicate a significant relationship between the Morningstar rating and the subsequent geometric mean. In the 3-year and 5-year sample 2 out of 5 and 2 out of 3 respectively indicated a significant relationship. However, taking a look at the significant signs of the coefficients, we discover that in the one-year samples of 2005, 2009 and 2010, the three-year sample of 2009-2011 and the 5-year sample of 2007-2011 the coefficients of lower rated funds were positive and significantly different from the 5-star coefficient. For example in 2009 a fund that received a 1-star rating at the beginning of the year showed a significantly higher performance than any of the other rating grades in 2009. Only some of the evaluation periods, namely 2006, 2011, 2006-2008, 2005-2009, 2006-2010 show significant coefficients with negative signs. These are mostly the one- or two-star funds, indicating some predictive power for the lower rated funds. The R^2 values in the model indicate that only around 3% of the performance measure variance are explained by the variance of the ratings. However, the R^2 values differ over the years from 0.0156 in 2005 to 32.82 in 2006.

Sharpe Ratio

The results for the Sharpe ratio are outlined in table 15. The results are similar to those of the geometric mean. Only 2 out of 7 overall F-values in the 1-year evaluation period indicate a significant relationship between the rating and the subsequent performance. 3 out 5 for the three-year samples and 2 out of 3 for the five-year sample show significance. The sign of the significant coefficients is also mixed over all evaluation periods. Dummy variable coefficients with a significant positive sign appear in the 1-year samples of 2005, 2006, 2009, 2010 and in the 3-year samples of 2009-2011.

Sample	γ_0 (5 star)	γ_1 (1 star)	γ_2 (2 star)	γ_3 (3-star)	γ_4 (4-star)	R^2	F-stat
1-year							
2005 – 1year	0.00364 (1.18)	0.00356 (1.12)	0.00437 (1.39)	0.00628 (1.90)*	0.00417 (1.20)	0.0156	1.37
2006 – 1year°	0.02714 (36.31)***	-0.00551 (-2.23)***	-0.00720 (-5.83)***	-0.00221 (-2.70)***	-0.000313 (-0.38)***	0.3282	12.60***
2007 – 1year°	0.02143 (19.62)***	-0.00457 (-1.33)	-0.00428 (-2.83)***	-0.000818 (-0.62)	-0.00103 (-0.62)	0.0418	2.10*
2008 – 1year	-0.05175 (-14.12)***	-0.00288 (-0.56)	0.00376 (0.86)	0.00416 (1.02)	-0.000140 (-0.03)	0.0135	1.33
2009 – 1year°	0.01714 (5.97)***	0.01171 (2.98)***	0.00887 (2.52)**	0.00732 (2.46)*	0.00418 (1.42)	0.1541	5.37***
2010 – 1year°	0.00686 (3.21)***	-0.00523 (-0.67)	0.00439 (1.85)*	0.00335 (1.46)	0.00353 (1.45)	0.0587	2.40*
2011 – 1year°	-0.01307 (-7.67)***	-0.00630 (-2.70)***	-0.00415 (-2.03)**	-0.00449 (-2.37)**	-0.00260 (-1.36)	0.0655	2.51**
3-years							
2005 – 3years	0.01674 (7.24)***	0.000100 (0.42)	0.00155 (0.66)	0.000813 (0.34)	0.00119 (0.48)	0.0203	0.46
2006 – 3years	-0.00134 (-1.80)*	-0.00207 (-1.14)	-0.00766 (-3.38)***	-0.00028 (-0.24)	-0.000439 (-0.48)	0.08	3.14**
2007 – 3years	-0.00242 (-1.80)*	-0.00256 (-1.51)	0.000530 (0.34)	0.000013 (0.09)	0.000440 (0.28)	0.0302	1.75
2008 – 3years	-0.00769 (-4.78)***	0.000993 (0.44)	0.00253 (1.33)	0.00231 (1.29)	0.00160 (0.88)	0.0259	0.62
2009 - 3years	0.00237 (1.37)	0.00598 (1.48)	0.00410 (2.09)**	0.00348 (1.94)*	0.00204 (1.14)	0.0593	2.51**
5years							
2005 – 5years	0.01003 (7.03)***	-0.00397 (-2.56)**	-0.00238 (-1.58)	-0.00163 (-1.05)	-0.00186 (-1.08)	0.0736	2.85**
2006 – 5years	0.00656 (8.27)***	-0.00149 (-1.00)	-0.00416 (-2.80)***	-0.00106 (-1.15)	-0.000945 (-1.09)	0.0409	2.02*
2007 – 5years	-0.00177 (-0.84)	0.00372 (1.33)	0.00422 (1.70)*	0.000840 (0.36)	0.00196 (0.77)	0.0185	1.47
* = indicates significance at the 10% level ** = indicates significance at the 5% level							
*** = indicates significance at the 1% level ° = indicates heteroscedasticity at least at 10% level							

Table 14: Results from Morningstar using Geometric Mean

Sample	γ_0 (5 star)	γ_1 (1 star)	γ_2 (2 star)	γ_3 (3-star)	γ_4 (4-star)	R^2	F-stat
1year							
2005 – 1year°	0.07017 (4.58)***	0.07800 (3.88)***	0.10633 (5.31)***	0.14719 (3.25)***	0.10250 (3.27)***	0.0340	1.82
2006 – 1year°	0.58710 (22.94)***	-0.09128 (-0.95)	-0.11290 (-1.13)	0.09291 (2.98)**	0.07884 (2.75)***	0.1906	6.59***
2007 – 1year°	0.47104 (13.54)***	-0.09889 (-1.08)	-0.08139 (-1.80)*	0.00224 (0.06)	-0.00945 (-0.19)	0.0235	1.61
2008 – 1year	-0.43192 (-21.07)**	-0.00266 (-0.09)	0.01287 (0.53)	0.01667 (0.73)	0.02278 (0.98)	0.0193	0.46
2009 – 1year°	0.21727 (7.82)***	0.10957 (2.83)***	0.09533 (2.64)***	0.07786 (2.70)***	0.04227 (1.48)	0.1413	4.95***
2010 – 1year	0.11323 (4.16)***	-0.04305 (-0.46)	0.006214 (2.03)**	0.04872 (1.63)	0.04675 (1.52)	0.0375	1.88
2011 – 1year	-0.15201 (-7.52)***	-0.04713 (-1.70)*	-0.02153 (-0.89)	-0.00954 (-0.42)	-0.000666 (-0.03)	0.0180	1.4
3years							
2005 – 3years	0.37929 (5.74)***	0.04724 (0.70)	0.04928 (0.74)	0.00928 (0.14)	0.01754 (0.25)	0.0149	1.35
2006 – 3years°	-0.01210 (-2.14)**	-0.03610 (-1.81)*	-0.08898 (-2.48)**	-0.01297 (-1.34)	-0.00435 (-0.64)	0.1738	6.10***
2007 – 3years	0.00444 (0.98)*	-0.03235 (-1.78)*	-0.00614 (-0.80)	-0.00327 (-0.55)	-0.00226 (-0.31)	0.0457	2.16*
2008 – 3years	-0.03742 (-2.81)***	0.00654 (0.35)	0.01817 (1.15)	0.01596 (1.07)	0.01744 (1.16)	0.49	0.0205
2009 - 3years°	0.05852 (2.83)***	0.07271 (1.57)	0.05076 (2.16)**	0.04394 (2.04)**	0.02157 (1.32)	0.0588	2.50**
5years							
2005 – 5years	0.04612 (2.99)***	0.03296 (1.96)*	0.03460 (2.13)**	0.04079 (2.43)**	0.02789 (1.50)	0.024	1.57
2006 – 5years°	0.09525 (20.32)***	-0.01938 (-2.00)**	-0.04937 (-1.77)*	-0.01219 (-1.98)*	-0.00951 (-1.63)	0.0775	3.02**
2007 – 5years°	-0.00240 (-0.53)	-0.03130 (-2.43)**	-0.00530 (-0.85)	-0.00474 (-0.82)	-0.00377 (-0.61)	0.0850	3.32*

* = indicates significance at the 10% level   ** = indicates significance at the 5% level

*** = indicates significance at the 1% level   ° = indicates heteroscedasticity at least at 10% level

Table 15: Results from Morningstar using Sharpe Ratio

8.3.2 Feri

The regression results for Feri are outlined in table 16 and 17. In the following we will look at the different results from the geometric mean and the Sharpe ratio as performance measures.

Geometric Mean

Results for Feri show much lower significance over all examination windows. Only 2 out of 7 of the 1-year samples indicate a significant relationship based on the F-value. Furthermore, only 2 out of 5 of the 3-year-samples and 1 out of 3 of the 5-year samples show a F-value that indicates significance. In general, the significant coefficients indicate a negative relationship for the lower-rated funds (1- and 2-star rating). No coefficient for the 4-star rated funds is significant, indicating that there is no difference in the performance of a 5-or 4-star rated fund.

Sharpe Ratio

The results for Feri using the Sharpe Ratio as a performance indicator do not differ substantially from those using the geometric mean. Only 1 out of 7 of the 1-year samples indicate a significant relationship based on the F-value. Furthermore, only 2 out of 5 of the 3-year-samples and 1 out of 3 of the 5-year samples show a F-value that indicates significance. However, most coefficients are negative, with the exception of the 1-year sample of 2008.

Sample	γ_0 (5-star)	γ_1 (1-star)	γ_2 (2-star)	γ_3 (3-star)	γ_4 (4-star)	R^2	F-stat
1-year							
2005 – 1year	0.00882 (12.56)***	-0.00196 (-1.76)*	-0.00169 (-1.88)*	-0.00103 (-1.14)	-0.00132 (-1.52)	0.0100	1.15
2006 – 1year	0.02751 (32.98)***	-0.00210 (-1.96)*	-0.00210 (-2.90)***	-0.00253 (-2.17)**	-0.00149 (-1.37)	0.0653	2.1*
2007 – 1year	0.02140 (15.28)***	-0.00035 (-0.15)	-0.00045 (-0.26)	0.0000522 (0.32)	0.00216 (0.71)	0.0264	0.38
2008 – 1year	-0.05051 (-14.23)	0.01200 (2.17)**	0.000058 (0.01)	-0.00231 (-0.54)	0.000767 (0.15)	0.0713	2.19*
2009 – 1year	0.02137 (12.69)	0.00677 (2.42)**	0.00135 (0.68)	0.0004848 (0.24)	0.000789 (0.39)	0.0522	1.83
2010 – 1year	0.01024 (5.17)	-0.0030 (-1.07)	-0.00205) (-0.96)	-0.00106 (-0.48)	-0.00126 (-0.58)	0.0378	0.48
2011 – 1year	-0.01664 (-8.27)	-0.00292 (-1.15)	-0.000924 (-0.41)	-0.002 (-0.94)	-0.00049 (-0.21)	0.0589	0.73
3-years							
2005 – 3years	0.01913 (35.00)***	-0.00098 (-1.39)	-0.00132 (2.15)**	-0.0000454 (-0.73)	-0.000942 (-01.05)	0.0528	0.77
2006 – 3years	-0.00166 (-1.11)	-0.00008 (-0.40)	-0.00155 (-0.89)	0.00166 (0.94)	0.00153 (0.86)	0.0622	2.04*
2007 – 3years	-0.00251 (-1.83)	0.00156 (0.66)	-0.000305 (-0.18)	-0.0003809 (-0.24)	0.000993 (0.57)	0.0302	0.44
2008 – 3years	-0.00510 (-3.68)	0.00202 (0.94)	-0.00233 (-1.31)	-0.00305 (-1.82)*	-0.00166 (-0.83)	0.0740	2.24*
2009 - 3years	0.00435 (5.13)	0.00177 (1.26)	-0.00098 (-0.98)	-0.000673 (-0.66)	0.000157 (0.16)	0.0374	1.58
5years							
2005 – 5years	0.00554 (15.49)***	-0.00179 (-1.54)	-0.000174 (-0.32)	0.0004912 (1.05)	0.0001307 (0.29)	0.0919	2.49*
2006 – 5years	0.00568 (5.86)***	-0.00082 (-0.57)	-0.00120 (-1.05)	0.000361 (0.31)	0.000298 (0.26)	0.0083	1.13
2007 – 5years	-0.00303 (-3.07)	0.00119 (0.70)	-0.00055 (-0.45)	-0.00097 (-0.84)	0.000466 (0.37)	0.0595	0.89

* = indicates significance at the 10% level   ** = indicates significance at the 5% level

*** = indicates significance at the 1% level   ° = indicates heteroscedasticity at least at 10% level

Table 16: Results from Feri using Geometric Mean

Sample	γ_0 (5-star)	γ_1 (1-star)	γ_2 (2-star)	γ_3 (3-star)	γ_4 (4-star)	R^2	F-stat
1-year							
2005 – 1year°	0.18796 (10.22)***	-0.05423 (-1.86)*	-0.04635 (-1.97)*	-0.03367 (-1.43)	-0.03924 (-1.72)*	0.0174	1.26
2006 – 1year°	0.065033 (21.69)***	0.06850 (1.66)	0.04004 (1.21)	0.04353 (1.31)	0.03788 (1.13)	0.0478	0.74
2007 – 1year	0.52891 (14.32)***	0.03631 (0.57)	0.00336 (0.07)	0.01821 (0.42)	0.01881 (0.40)	0.0096	0.14
2008 – 1year	-0.40062 (-24.23)	0.04994 (1.94)*	-0.01691 (-0.79)	-0.00109 (-0.06)	0.00109 (0.05)	0.0598	1.99*
2009 – 1year	0.25904 (17.41)	0.03852 (1.56)	0.01729 (0.98)	0.000859 (0.05)	0.00262 (0.15)	0.0079	1.12
2010 – 1year	0.15554 (6.78)	-0.03635 (-1.12)	-0.02194 (-0.89)	-0.01289 (-0.51)	-0.014 (-0.58)	0.0345	0.44
2011 – 1year	-0.1355 (-5.33)	-0.02495 (-0.78)	-0.01102 (-0.39)	-0.002556 (-0.95)	-0.00618 (-0.21)	0.0455	0.56
3-years							
2005 – 3years	0.45578 (29.59)***	-0.00430 (-0.18)	-0.02051 (-1.04)	-0.00507 (-0.26)	-0.01056 (-0.55)	0.0246	0.35
2006 – 3years	-0.00368 (-0.29)	-0.01172 (-0.63)	-0.01238 (-0.82)	0.01521 (1.01)	0.01495 (0.98)	0.0762	2.3*
2007 – 3years	0.00371 (0.37)	0.00723 (0.42)	-0.00585 (-0.48)	-0.000312 (-0.03)	0.01117 (0.89)	0.0486	0.72
2008 – 3years	-0.01338 (-1.39)	0.00851 (0.57)	-0.02594 (-2.09)**	-0.02425 (-2.09)**	-0.01801 (-1.29)	0.0959	2.64**
2009 - 3years	0.08393 (9.12)	0.02030 (1.33)	-0.01062 (-0.98)	-0.00781 (0.70)	0.00195 (0.18)	0.0449	1.71
5years							
2005 – 5years°	0.08193 (18.97)***	-0.01956 (-1.50)	-0.00278 (-0.40)	0.00602 (1.07)	0.000928 (0.16)	0.0659	2.04*
2006 – 5years	0.08641 (9.56)***	-0.00874 (-0.66)	-0.00973 (-0.91)	0.00317 (0.7692)	0.00303 (0.28)	0.0621	0.98
2007 – 5years	-0.00599 (-0.78)	0.00431 (0.33)	-0.00835 (-0.88)	-0.00755 (-0.84)	0.00573 (0.59)	0.0093	1.14

* = indicates significance at the 10% level   ** = indicates significance at the 5% level

*** = indicates significance at the 1% level   ° = indicates heteroscedasticity at least at 10% level

Table 17: Results from Feri using Sharpe Ratio

8.3.3 Overview

Table 17 and 18 give a comprehensive overview over the signs and significance of the dummy variable coefficients. The tables clearly outline that the predictive power of the two raters is poor. While Morningstar exhibits limited predictive power for the performance measure of the geometric mean for the short-term periods of one year, the ratings show no predictive power for the performance measure of the Sharpe ratio in the short-term periods. For the long-run samples the Morningstar ratings show however some predictability for the performance measured by the Sharpe ratio.

While Morningstar exhibits little predictive power, Feri ratings clearly lack any predictability characteristics. Most dummy variable coefficients were not significant and switched between positive and negative signs.

Sample		eometric Mean	Sharpe Ratio
1-year	Positive	5	10
	Negative	8	3
	Not significant	15	16
3-year	Positive	2	2
	Negative	1	3
	Not significant	17	14
5-year	Positive	1	0
	Negative	2	4
	Not significant	9	7

Table 18: Overview over Morningstar Dummy Variable Coefficients

Sample		Geometric Mean	Sharpe Ratio
1-year	Positive	2	1
	Negative	5	3
	Not significant	21	24
3-year	Positive	-	-
	Negative	3	2
	Not significant	17	18
5-year	Positive	-	-
	Negative	-	-
	Not significant	12	12

Table 19: Overview over Feri Dummy Variable Coefficients

9 Conclusion

The results from the empirical analysis clearly indicate that the Morningstar rating shows very limited predictive power and that the Feri rating shows almost no predictive power for the German market. Different from Morey and Gottesman's (2006) findings for the US market, there is no improvement of the predictive power of the Morningstar ratings for German funds after the methodology change in 2006 in Europe. The question of *mutual fund ratings as a predictor for future performance* can therefore clearly be denied for the German market and no value is generated for the individual investor from a predictability feature of the ratings. Especially, in the case of Feri this is an interesting finding, since Feri claims to provide fund ratings that are able to predict future performance. Clearly, this statement does not resemble the truth and Feri is therefore misleading their investors.

When evaluating the initial problem statement of this thesis: *How valuable are ratings for an investor*, other aspects of the ratings than predictability have to be considered, because value for the individual investor might also stem from them. For this reason we outlined the specific methodologies of the three most important raters in Germany. As some raters claim their rating should rather be used as a source to screen possible investment funds, this study elaborated their individual strength and weaknesses in assessing the past features of the funds compared to their competitors. Each of the raters shows some weaknesses and strengths compared to the others. However, even when anticipating the fund rating as an indicator for past performance as suggested by rating providers the individual method used by the three raters is still questionable. While the different raters handle the outlined problems in a different manner, they all suggest a link between rating and performance. All information is comprised in one single number. What seems comprehensive at first sight gives the individual investor a wrong impression of the predictive value of a rating and hides the specific features of a fund from him. The information value of the ratings is therefore limited, as they do not present the implied and processed information of the individual features of the fund. Lipper is the only rating provider, which shows a more detailed overview over the characteristics of the fund. Their separated assessment of fund expenses, risk, return, and persistency gives a more detailed summary of the features of the fund, while Morningstar and Feri hide these features in their complex

methodology. However, within these categories Lipper does also not provide a detailed description of the rating input.

Given the outlined drawbacks one can conclude that the rise of the mutual fund rating industry clearly doesn't stem from its predictability. Furthermore, the information value for investors is also limited as raters do not provide the input data of funds for their rating, which might contain valuable information for the individual investor on a comprehensive basis. The predominant influence of Morningstar and co. therefore seems to be built on the misunderstanding or the irrationality of the investor.

10 Further Research

The thesis indicates further research on the field of mutual fund ratings. As the industry trends to provide an increasing number of qualitative fund ratings for investors, it is interesting to examine their predictive power. Especially, since Morningstar considers these ratings as a recommendation that provides actual predictability of future returns. As the ratings were first introduced in 2011, a first study should be commenced in 2014, analyzing the question of predictability.

Bibliography

Adkisson, J. A. and Fraser, D. R. (2004) *Is there still an Age Bias in the Morningstar Ratings?* Working Paper. A&M University Texas.

Alexander, G. J., Jones, J. D. and Nigro, P. J. (1998) *Mutual Fund Shareholder: Characteristics, Investor Knowledge and Sources of Information.* Financial Services Review. Vol. 7. pp. 301-316.

Amenc, N. and Le Sourd, V. (2007) *Rating the Ratings.* edhec-risk.com. retrieved from: http://faculty-research.edhec.com/servlet/com.univ.collaboratif.utils.LectureFichiergw?ID_FICHIER=1328885972064 on 23.06.2012.

Akerlof, G. A. (1970) *The Market for "Lemons": Quality Uncertainty and the Market Mechanism.* The Quarterly Journal of Economics. Vol. 84 (3). pp. 488-500.

Anderson, S. and Ahmed, P. (2005) *Mutual Funds: 50 Years of Research Findings.* Springer.

Barber, B. M., Odean, T. and Zheng, L. (2003) *Out of Sight, Out of Mind: The Effect of Expenses on Mutual Fund Flows*. The Journal of Business. Vol. 78 (6). pp. 2095-2120.

Barzel, Y. (1982) *Measurement Cost and the Organization of Markets.* Journal of Law and Economics. Vol. 5 (1). pp. 27-48.

Blake, C. R. and Morey, M. R. (2000) *Morningstar Ratings and Mutual Fund Performance.* Journal of Financial and Quantitative Analysis. Vol. 35 (3). pp. 451-483.

Bloomberg (2011) *Credit Rating Fees Rise Faster Than Inflation As Governments Downgraded.* Retrieved from: http://www.bloomberg.com/news/2011-11-15/credit-rating-fees-rise-faster-than-inflation-as-governments-fret-expenses.html on 12.04.2012.

Blume, M. E. (1998) *An Anatomy of Morningstar Ratings*. Financial Analysts Journal. Vol. 54 (2). pp. 19-27.

Brison, G. P., Singer, B. D. and Beebower G. L. (1991) *Determinants of Portfolio Performance II: An Update.* Financial Analysts Journal. Vol. 47 (3). pp. 40-48.

Brown, S. J. and Goetzmann, W. N. (1997) *Mutual Fund Styles*. Journal of Financial Economics. Vol. 43 (3). pp. 373-399.

Brown, S. J., Goetzmann, W. N., Ibbotson, R. G. and Ross, S. A. (1992) *Survivorship Bias in Performance Studies.* Review of Financial Studies. Vol. 5 (4). pp. 553-580.

BVI - Bundesverband Investment und Asset Management e.V. (2004) *Transparenz-Standards für die Beurteilung von Investmentfonds durch Ranking-/Ratingagenturen.* Retrieved from: http://www.bvi.de/de/bvi/leitlinien_standards/rrts/rrts_broschuere06/RRTS_Broschuere_2006.pdf on 24.06.2012.

BVI - Bundesverband Investment und Asset Management e.V. (2012) *Öffentlich zum Vertrieb zugelassene Fonds.* Retrieved from: http://www.bvi.de/de/statistikwelt/wirtschaft_und_investmentfonds/zum_vertrieb_zugel_fonds/index.html on 13.05.2012.

Capon, N., Fitzsimons, G. J. and Prince, R. A. (1996) *An Individual Level Analysis of the Mutual Fund Investment Decision.* Journal of Financial Services Research. Vol. 10 (1). pp. 59-82.

Carhart, M. M. (1997) *On Persistence in Mutual Fund Performance.* The Journal of Finance. Vol. 52 (1). pp. 57-82.

CESR - Committee Of European Securities Regulators (2010) *CESR's guidelines on the methodology for the calculation of the synthetic risk and reward indicator in the Key Investor Information Document.* Retrieved from: http://www.esma.europa.eu/system/files/10_673.pdf on 01.10.2012.

Chan, L. K. and Lakonishok, J. (2004) *Value and Growth Investing: Review and Update. Financial Analysts Journal*. Vol. 60, (1). pp. 71-86.

Chunhachinda, P., Dandapani, K., Hamid, S. and Prakash, A. J. (1997) Portfolio selection and skewness: Evidence from international stock markets. Journal of Banking & Finance. Vol. 21 (2). pp. 143–167.

Clark, A. (2003) *The Use of Hurst and Effective Return in Investing*. Quantitative Finance. Vol. 5 (1). pp. 1-8.

Crouhy, M.G., Jarrow, R. A. and Turnbull, S. M. (2008) *The Subprime Credit Crisis of 2007.* Journal of Derivatives. Vol. 16 (1). pp. 81-110.

Damato, K. (1996) *Morningstar Edges Toward One-Year Ratings.* Wall Street Journal. April 5.

Del Guercio, D. and Tkac, P. (2008) *Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow.* Journal of Financial and Quantitative Analysis. Vol. 43 (4). pp. 907-936.

diBartolomeo, D. and Witkowski, E. (1997) *Mutual Fund Misclassification: Evidence Based on Style Analysis.* Financial Analysts Journal. Vol. 53 (5). pp. 32-43.

Dacorogna, M. M., Gençay, R., Müller, U. A. and Pictet, O. V. (2001) *Effective Return, Risk Aversion and Drawdowns.* Physica A. Vol. 289 (1-2). pp. 229-248.

Duret, A.-S., Hereil, P., Mitaine, P., Moussavi, N. and Roncalli, T. (2008) *Fund Rating Systems and Performance Predictability*. Working Paper. Université d'Évry.

EFAMA - European Fund and Asset Management Association (2011) *Fund Fees In Europe: Analyzing Investment Management Fees, Distribution Fees, And Operating Expenses.* Retrieved from: http://www.efama.org/index.php?option=com_docman&task=cat_view&gid=506&Itemid=-99 on 24.06.2012.

Eisenhardt, K. M. (1985) *Control: Organizational and Economic Approaches*. Journal of Management Science. Vol. 31 (2). pp. 134-149.

Elton, E. J., Gruber, M. J. and Blake, C. R. (1995) *The Persistence of Risk-Adjusted Mutual Fund Performance.* NYU Working Paper.

Fama, E.F. (1965) *The Behavior of Stock-market Prices.* Journal of Business. Vol. 38. pp. 34-105.

Fama, E. F. (1970) *Efficient Capital Markets: A Review of Theory and Empirical Work.* Journal of Finance. Vol. 25 (2). pp. 383-417.

Feri EuroRating Services AG (Feri) (2012) *Das Feri Fondrating.* Retrieved from:
http://www.feri.de/Content/Frr/Files/Fondsplattform/PUBLICARCHIVE/Fondsrating.pdf on 01.04.2012.

Feri EuroRating Services AG (Feri) (2012a) *Feri Fonds Portal.* Retrieved from: http://www.feri-fund-ra-ting.com/Default.aspx?Name=FundsRatingGermany&Content=TopFonds&Lang=en

on 31.10.2012.

Fernando, D., Klapper, L., Sulla, V. and Vittas, D. (2003) *The Global Growth of Mutual Funds.* World Bank Policy Research Working Paper 3055. The World Bank.

Fink, M. P. (2011) *The Rise of Mutual Funds: An Insider's View.* Second Edition. OUP USA.

Fischhoff, B., Slovic, P. and Lichtenstein, S. (1977) *Knowing with certainty: The appropriateness of extreme confidence.* Journal of Experimental Psychology 3. pp.552-564.

Füss, R., Hille, J., Rindler, P., Schmidt, J. and Schmidt, M. (2010). *From Rising Stars and Falling Angels: On the Relationship Between the Performance and Ratings of German Mutual Funds.* The Journal of Wealth Management. Vol. 13 (1). pp. 75-90.

Ft.com (2012) *Feri offers UK rating.* Retrieved from: http://www.ft.com/intl/cms/s/0/93f4bc2c-9103-11df-b297-00144feab49a.html#axzz1rGAo43f6 on 23.06.2012.

Gerrans, P. (2004) *Australian Managed Fund Ratings and Individual Investors.* Australian Journal of Management. Vol. 29 (1). pp. 87-107.

Gerrans, P. (2006) *Morningstar Ratings and Future Performance.* Accounting and Finance. Vol. 46 (4). pp. 605-628.

Goldberg, L. R., Hayes, M., Menchero, J. and Mitra, I. (2009) Extreme Risk Analysis. Journal of Performance Management. Vol. 14 (3). pp. 17–30.

Hendricks, D., Patel, J. and Zeckhauser, R. (1993) *Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance,* 1974-1988. T*he Journal of*

Finance. Vol. 48 (1). pp. 93-130.

Hereil, P., Mitaine, P., Moussavi, N. and Roncalli, T. (2010) *Mutual Fund Ratings and Performance Persistence.* White Paper, Quant Research by Lyxor.

Herzog, D. A. (2007) *Math You Can Really Use – Every Day*. Wiley, Publishing Inc.

Hill, C. A. (2002) *Rating Agencies Behaving Badly: The Case of Enron.* Connecticut. Law Review. Vol. 35. pp. 1145-1155.

Huang, J., Sialm, C. and Zhang, H. (2008) *Risk Shifting and Mutual Fund Performance.* Review of Financial Studies. Vol. 24 (8). pp. 2575-2616.

ICI - Investment Company Institute (2002) *U.S. Household Ownership of Mutual Funds in 2002: Fundamentals.* Investment Company Institute Research. Retrieved from: http://www.ici.org/pdf/fm-v11n5.pdf on 23.06.2012.

ICI - Investment Company Institute (2011) 2011 *Investment Company Fact Book.* 51st Edition. Retrieved from http://www.ici.org/pdf/2011_factbook.pdf on 01.06.2012.

ICI - Investment Company Institute (2012a) *2012 Investment Company Fact Book.* 52nd Edition. Retrieved from http://www.ici.org/pdf/2012_factbook.pdf on 01.05.2012.

ICI - Investment Company Institute (2012) *A Guide to Mutual Funds.* Retrieved from http://www.ici.org/investor_ed/brochures/bro_g2_mfs#_Variety on 01.05.2012.

Jensen, A. C. (1968) *The Performance of Mutual Funds in the Period 1945-1964. The Journal of Finance.* Vol. 23 (2). pp. 389-416.

Jones, C. P. (2009) *Investments: Analysis and Management.* Eleventh Edition. John Wiley & Sons.

Jones, M. A. and Smythe, T. (2003*) The Information Content of Mutual Fund Print Advertising.* The Journal of Consumer Affairs. Vol. 37 (1). pp. 22-41.

Khorana, A., and Servaes, H. (2012) *What Drives Market Share in the Mutual Fund Industry?* Review of Finance. Vol. 16 (1). pp. 81-113.

Khorana, A. and Nelling, E. (1998) *The Determinants and Predictive Ability of Mutual Fund Ratings.* The Journal of Investing. Vol. 7 (3). pp. 61-66.

Kim, M., Shukla, R. and Tomas, M (1995) *Wolf in Sheep's Clothing: Do Mutual Fund Objectives Tell the Truth?* Working paper. Syracuse University.

Kräussl, R. and Sandelowsky R. M. (2007) *The Predictive Performance of Morningstar's Mutual Fund Ratings.* Working Paper. VU University Amsterdam.

Kräussl, R. (2007) *Improvement and Limitation of the Revised Morningstar Fund Rating Methodology.* in: Gregoriou, G. N.: Diversification and Portfolio Management of Mutual Funds. Palgrave McMillan. pp. 121-137.

Langohr, H. and Langohr, P. (2009) *The Rating Agencies and Their Credit Ratings: What They Are, How They Work and Why They are Relevant.* John Wiley & Sons.

Levich, R. M., Majnoni, G. and Reinhart, C. (2002) *Ratings, Rating Agencies and the Global Financial System.* Springer.

Lipper (2008) *Lipper Global Classification – Extended: Definition Document.* Retrieved   from:

http://www.lipperweb.com/docs/support/DataChange/Extended_LGC_Definitions.pdf on 01.11.2012.

Lipper (2011) Lipper Leaders – *Rating System US Methodology.* Retrieved from: http://www.lipperweb.com/docs/research/leaders/LLMethodologyUS_V3.pdf on 01.08.2012.

Lipper (2012) Lipper Leaders – *Rating System International Methodology*. Retrieved from: http://www.lipperweb.com/docs/AboutUs/LLMethodology_Intl_V2_A4.pdf on 01.08.2012.

Lipper (2012a) N*umber of Funds in German Fund Universe.* Retrieved from http://www.lipperleaders.com/index.aspx on 07.05.2012.

Lipper (2012b*) Lipper Content for Marketing and Advertising - Say something new to your Investors. Retrieved from:*
http://www.lipperweb.com/docs/AboutUs/Marketing_Sample_Booklet_WEB.pdf on 01.10.2012.

Lisi, F. (2010) *The Weight of Risk in the Morningstar Rating for Mutual Funds*. Working Paper. University of Padua.

Mandelbrot, B. (1963) *The Variation of Certain Speculative Prices.* Journal of Business. Vol. 36. pp. 394-419.

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

Massa, M. and Patgiri, R. (2009) *Incentives and Mutual Fund Performance: Higher Performance or Just Higher Risk Taking?* Review of Financial Studies. Vol. 22 (5). pp. 1777-1815.

McNeil (1999) *Extreme Value Theory for Risk Managers.* Department of Mathematics. ETH Zürich. Working Paper.

MLP (2011) *Press Release: MLP übernimmt plangemäß die ausstehenden Anteile an Feri.* Retrieved from: http://www.mlp-ag.de/#/presse/pressemitteilungen/2011/110420-pm/ on 23.06.2012.

MLP (2012) *MLP Fact Sheet.* Retrieved from: http://www.mlp-ag.de/#/unternehmensprofil/portraet/factsheet on 23.06.2012.

Modigliani, F. and Modigliani, L. (1997) *Risk-Adjusted Performance.* The Journal of Portfolio Management. Vol. 23 (2). pp. 45-54.

Moisson, E. for Lipper (2012) *Beating the Benchmark*. Retrieved from: http://funds.uk.reuters.com/UK/pdf.asp?language=UNK&dockey=1523-4357-1MUR4NKAQL1P7BS2U69G44FT7O on 23.06.2012.

Moody's (2012) *Ratings Policy and Approach.* Retrieved from: http://www.moodys.com/ratings-process/Ratings-Policy-Approach/002003.

Morey, M. R. and Gottesman, A. (2006) *Morningstar Mutual Funds Rating Redux.* Journal of Investment Consulting. Vol. 8 (1). pp. 25-37.

Morey M.R. and Blake, C. R. (2000) *Morningstar Ratings and Mutual Fund Performance.* Journal of Financial and Quantitative Analysis. Vol. 35 (3). pp. 451-483.

Morey, M. R. (2000) *The Morningstar Mutual Fund Ratings: What Investors Should Know.* Research Dialogue. TIAA-CREF Institute. Vol. 64. Retrieved from: http://www.tiaa-crefinstitute.org/institute/research/dialogue/rd_64.html on 24.04.2012.

Morey, M. R. (2002) *Mutual Fund Age and Morningstar Ratings.* Financial Analysts Journal. Vol. 58 (2). pp. 56-63.

Morningstar (2004) *Fact Sheet: The Morningstar Style Box*. Retrieved from: http://corporate.morningstar.com/uk/documents/MethodologyDocuments/FactSheets/ EU_MorningstarStyleBox_FactSheet.pdf on 23.06.2012.

Morningstar (2005) Rating the Star Rating. Received upon request on 12.04.2012.

Morningstar (2005a) *Morningstar Europe Rating - Studie: Was sagen die Sterne?* Retrieved from: http://www.fondsprofessionell.de/upload/attach/504808.pdf on 01.05.2012.

Morningstar (2009) *The Morningstar Rating Methodology.* Retrieved from: http://corporate.morningstar.com/de/documents/MethodologyDocuments/Methodolog yPapers/MorningstarFundRating_Methodology.pdf on 23.06.2012.

Morningstar (2010*) The Morningstar Rating for Funds.* Retrieved from: http://news.morningstareurope.com/news/im/msuk/PDFs/Morningstar-Rating-Fact-Sheet1.pdf on 01.04.2012.

Morningstar (2010a) *How Expense Ratios and Star Ratings Predict Success*. Retrieved from: http://news.morningstar.com/articlenet/article.aspx?id=347327 on 23.06.2012.

Morningstar (2012) *Morningstar Analyst Rating for Mutual Funds.* Retrieved from: http://www.morningstar.com/InvGlossary/morningstar-analyst-rating-for-funds.aspx on 02.04.2012.

Morningstar (2012a) *Morningstar Company Site.* Retrieved from: http://corporate.morningstar.com/au/asp/subject.aspx?xmlfile=5679.xml on 03.05.2012.

Morningstar (2012b) *Morningstar Fundscreener.* Retrieved from: http://tools.morningstar.de/de/fundscreener/default.aspx?Site=de&LanguageId=de-DE on 03.11.2012.

Park, H. M. (2002) Using Dummy Variables in Regression. Indiana University at Bloomington. Working Paper.

Pozen, R. C. (1998) *The Mutual Fund Business.* The MIT Press.

Qian, B., and Rasheed, K. (2004) *Hurst exponent and financial market pre-dictability.* IASTED conference on "Financial Engineering and Applications". pp. 203–209.

Rachev, S., Martin, R. D., Racheva, B. and Stoyanov, S. (2006) *Stable ETL Optimal Portfolios and Extreme Risk Management,* in *Risk Assessment: Contribution to Economics.* pp. 235-263. Springer.

Reiter, C. (2007) *Fractals, Visualization and J.* Third Edition. Lulu.

Risager, O. (2009) *Investing in Value Stocks.* McGraw-Hill.

Sharpe, W. F. (1975) *Adjusting for Risk in Portfolio Performance Measure-ment.* The Journal of Portfolio Management. Vol. 1 (2). pp. 29-34.

Sharpe, W. F. (1998) *Morningstar's Risk-Adjusted Ratings.* Financial Analysts Journal. Vol. 54 (4). pp. 21-33.

Smith, R.C., Walter, I. and DeLong, G. (2012) *Global Banking.* Third Edition. OUP USA.

Spritzer, F. (2001) *Principles of Random Walk.* Second Edition. Springer-Verlag.

Starks, S. T. (1987) *Performance Incentive Fees: An Agency Theoretic Approach.* Journal of Financial and Quantitative Analysis. Vol. 22 (1). pp 17-32.

Stiglitz, J. E. (2000) *The Contribution of the Economics of Information to Twentieth Century Economics.* The Quarterly Journal of Economics. Vol. 115 (4). pp. 1441-1478.

Stutzer, M. (2005) *Mutual Fund Ratings: What is the Risk in Risk-Adjusted Fund Returns?* Working Paper. University of Colorado.

S&P (2012) *Fund Research.* Retrieved from: http://www.standardandpoors.com/funds/main/en/eu on 25.04.2012.

SEC - United States Securities and Exchange Commission (2003) *Report on the Role and Function of Credit Rating Agencies in the Operation of the Securities Markets As Required by Section 702(b) of the Sarbanes-Oxley Act of 2002.* Washington, D.C.

Teverovsky, V., Taqqu, M. S. and Willinger, W. (1999) *A critical look at Lo's modified R/S statistic.* Journal of Statistical Planning and Inference 80. pp. 211–227.

Vaihekoski, M. (2009) *On the calculation of the risk free rate for tests of asset pricing models.* Retrieved from SSRN: http://ssrn.com/abstract=958471 on 01.09.2012.

Van Harlow, W., Brown, K. C. and Zhang, H. (2011) *Staying the Course: The*

Role of Investment Style Consistency in the Performance of Mutual Funds. Working Paper. University of Texas.

Wall Street Journal Homepage (2012) *Mutual Fund Scorecards.* Retrieved from: http://online.wsj.com/fund/page/fund_scorecards.html on 23.06.2012.

Watson, J., Wickramanayke, J. and Premachandra, I. M. (2011) *The Value of Morningstar Ratings: Evidence using Stochastic Data Envelopment Analysis.* Managerial Finance. Vol. 37 (2). pp. 94-116.

Weingärtner, A. (2009) *Performance-Analyse & Fondbewertungssysteme: Eine empirische Untersuchung der S&P Fund Stars und des S&P Fund Management Ratings.* Books on Demand GmbH.

White, H, (1980) *A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity.* Econometrica. Vol. 48 (4). pp. 817-838.

Wittrock, C. (2003) *Integration investorspezifischer Belange bei der Fondsauswahl,* in: Achleitner A./Everling O. (publisher): *Fondsrating: Qualitätsmessung auf dem Prüfstand - Verfahren, Kriterien und Nutzen.* Gabler Verlag. pp. 3-16.

World Federation of Exchanges (2012) *Number of listed companies.* Retrieved from http://www.world-exchanges.org/statistics/annual/2010/equity-markets/number-listed-companies on 01.05.2012.

Appendix 1: Morningstar's Analyst Rating™

In November 2011 Morningstar introduced the Analyst Rating™ as a qualitative and forward-looking counterpart to the star rating. The rating is based on five pillars:

(1) **Fund Management:** The Portfolio Manager will be assessed based on his or her experience, knowledge and the overall work load, but also on the support from analysts and the incentive scheme.

(2) **Fund Company:** In order to assess the fund's future prospects Morningstar thinks it is of great interest to evaluate the investment firm as a whole. Interesting factors are the ownership structure, stability of management and the compensation structure, as well as the business culture

(3) **Process:** A main part of the analysis will be based on the investment decision. Not only the stock picking process will be assessed but also the portfolio construction. The process will also assess the risks that are inherent with the investment style. The analysts do take historical but also forward looking indicators and ratios into consideration when evaluating the investment process.

(4) **Performance**: The performance evaluation is not only based on past indicators but aims to provide forward-looking indication. In order to do so Morningstar looks through the performance of a whole business cycle.

(5) **Costs:** Morningstar beliefs that there is a high connection between future performance and the costs of a fund. The ratio that is most important in this discussion is the TER (Total Expense Ratio) of the fund.

Investment Analysts assess these factors of a fund according to his peer group.

The ratings can be positive, neutral or negative:

(1) Positive: There are three categories of positive ratings

Gold: This is the highest rating a fund can achieve. These funds score above average in all fund categories.

Silver: This is the second highest rating a fund can achieve. These fund do not have above average scores in all 5 categories but in most of them

Bronze: This is the lowest positive rating a fund can achieve. In this category funds show still more advantage than disadvantages in the 5 categories outlined above.

(2) Neutral: This category is assigned to funds that do neither outperform nor underperform.

(3) Negative: Funds in this category show deficits according to their peer group. The analysts are convinced that these funds will fall behind its competitors' performance in the long run.

Whether a fund gets rated depends on its volume, quality and investors' demand. Contrary to the star rating the rating there is no fixed percentage of funds assigned to each rank. The decision to rank a fund as positive, negative or neutral will solely depend on the analysts' opinion. The revision of the rating decision will only be assessed on a yearly basis. Funds that have a higher volume or a higher rating will be reviewed more frequently.

Appendix 2: Comparison of the Quantitative and Qualitative Morningstar Rating

In the light of the two provided ratings of Morningstar it is interesting to evaluate the differences between the two ratings. The following section will therefore outline information of the two ratings for all mutual funds on the German Morningstar Database that have both the qualitative and the quantitative rating. All 398 ratings were published between 10.02.2011 and 29.03.2012.

Table X shows how often each qualitative category was assigned to a fund.

Qualitative Fund Rating	Number of Funds
Gold	26
Silver	81
Bronze	115
Neutral	125
Negative	14

Table 20: Number of funds in each Morningstar Analyst Rating

It is important to note that there is an upwards bias in the ratings. 222 funds received a positive rating and only 125 a neutral, and 14 a negative rating. According to information from Morningstar this bias is due to the fact that the new qualitative rating was first conducted for those funds that show a high demand from investors which are often the best performing ones. As Morningstar wants to increase the number of rated funds from 400 to 1500 at the end of 2012, also less popular and therefore maybe worse performing funds will get a qualitative rating. This might increase the number of negative and neutral ratings.

When assessing the 1-and 2-rated funds we can find, that 15 out of 54 funds received a positive qualitative rating, i.e. the analyst thinks that the overall advantages overweigh the disadvantages of the fund. For the 3-star funds more than 70 out of 125 received a positive rating. Investors that only invest in 4-or 5-star funds therefore let good investment opportunities pass.

Appendix 3: Morningstar Rating Categories in Germany[10]

Aktien Afrika	Aktien weltweit Nebenwerte	Branchen: Immobilienaktien Europa
Aktien Afrika & Naher Osten	Aktien weltweit Standardwerte Blend	Branchen: Immobilienaktien Global
Aktien Afrika & Naher Osten Sonstige	Aktien weltweit Standardwerte Growth	Branchen: Immobilienaktien Nordamerika
Aktien Amerika Sonstige	Aktien weltweit Standardwerte Value	Branchen: Industriematerialien
Aktien ASEAN	Alt - Aktien Market Neutral	Branchen: Infrastruktur
Aktien Asien ohne Japan	Alt - Dachfonds - Aktien	Branchen: Kommunikation
Aktien Asien Sonstige	Alt - Dachfonds - Multistrategy	Branchen: Konsumgüter und -dienstleistungen
Aktien Asien-Pazifik mit Japan	Alt - Dachfonds - Sonstige	Branchen: Ökologie
Aktien Asien-Pazifik ohne Japan	Alt - Debt Arbitrage	Branchen: Private Equity
Aktien Australien und Neuseeland	Alt - Diversified Arbitrage	Branchen: Rohstoffe
Aktien Belgien	Alt - Event Driven	Branchen: Technologie

[10] Morningstar (2012b) *Morningstar Fundscreener.* Retrieved from:

http://tools.morningstar.de/de/fundscreener/default.aspx?Site=de&LanguageId=de-DE on 03.11.2012.

Aktien Brasilien	Alt - Global Macro	Branchen: Versorger
Aktien BRIC	Alt - Long/Short Aktien Europa	Branchen: Wasser
Aktien China	Alt - Long/Short Aktien Großbritannien	Euro Money Market PP
Aktien Dänemark	Alt - Long/Short Aktien Schwellenländer	Französisch PEA Eonia SWAP
Aktien Deutschland	Alt - Long/Short Aktien USA	Garantiefonds
Aktien Deutschland Nebenwerte	Alt - Long/Short Aktien weltweit	Geldmarkt AUD
Aktien EMEA	Alt - Long/Short Debt	Geldmarkt CAD
Aktien Euroland flexibel	Alt - Multistrategy	Geldmarkt CHF
Aktien Euroland mittelgroß	Alt - Systematic Futures	Geldmarkt DKK
Aktien Euroland Nebenwerte	Alt - Volatilität	Geldmarkt EUR
Aktien Euroland Standardwerte	Alt - Währungen	Geldmarkt EUR Kurzfristig
Aktien Europa außer Großbritannien Nebenwerte	Anleihen Asien	Geldmarkt GBP
Aktien Europa außer Großbritannien Standardwerte	Anleihen AUD	Geldmarkt GBP Kurzfristig
Aktien Europa flexibel	Anleihen CAD	Geldmarkt NOK
Aktien Europa mittelgroß	Anleihen CHF	Geldmarkt SEK
Aktien Europa Nebenwerte	Anleihen CHF Kurzläufer	Geldmarkt SGD
Aktien Europa Sonstige	Anleihen Dänemark	Geldmarkt Sonstige
Aktien Europa Standardwerte Blend	Anleihen DKK Langläufer	Geldmarkt USD
Aktien Europa Standardwerte Growth	Anleihen EUR diversifiziert	Geldmarkt USD Kurzfristig
Aktien Europa Standardwerte Value	Anleihen EUR diversifiziert Kurzläufer	Immobilienaktien Sonstige
Aktien Finnland	Anleihen EUR flexibel	Immobilienfonds Euroland
Aktien Frankreich Nebenwerte	Anleihen EUR hochverzinslich	Immobilienfonds Europa
Aktien Frankreich Standardwerte	Anleihen EUR inflationsgesichert	Immobilienfonds Global
Aktien Global Frontier Markt	Anleihen EUR Langläufer	Immobilienfonds Sonstige
Aktien Global währungsgesichert	Anleihen EUR ultra-short	Kapitalschutz
Aktien Greater China	Anleihen Europa	Kategorie
Aktien Griechenland	Anleihen Europa hochverzinslich	Laufzeitfonds 2000-2014
Aktien Großbritannien Ertrag	Anleihen GBP diversifiziert	Laufzeitfonds 2015+

Aktien Großbritannien flexibel	Anleihen GBP diversifiziert Kurz-läufer	Mischfonds Asien
Aktien Großbritannien mittelgroß	Anleihen GBP flexibel	Mischfonds CHF aggressiv
Aktien Großbritannien Nebenwer-te	Anleihen GBP hochverzinslich	Mischfonds CHF ausgewogen
Aktien Großbritannien Standard-werte Blend	Anleihen GBP inflationsgesichert	Mischfonds CHF defensiv
Aktien Großbritannien Standard-werte Growth	Anleihen Global	Mischfonds Emerging Markets
Aktien Großbritannien Standard-werte Value	Anleihen Global CHF	Mischfonds EUR aggressiv
Aktien Hong Kong	Anleihen Global CHF-hedged	Mischfonds EUR ausgewogen
Aktien Indien	Anleihen Global EUR	Mischfonds EUR defensiv
Aktien Indonesien	Anleihen Global EUR-hedged	Mischfonds EUR flexibel
Aktien Italien	Anleihen Global GBP-hedged	Mischfonds GBP aggressiv
Aktien Japan Nebenwerte	Anleihen Global hochverzinslich	Mischfonds GBP ausgewogen
Aktien Japan Standardwerte	Anleihen Global hochverzinslich EUR-hedged	Mischfonds GBP defensiv
Aktien Kanada	Anleihen Global hochverzinslich GBP-hedged	Mischfonds GBP flexibel
Aktien Korea	Anleihen Global ILS	Mischfonds NOK
Aktien Lateinamerika	Anleihen Global Sonstige (hedged)	Mischfonds SEK defensiv
Aktien Malaysien	Anleihen Global USD	Mischfonds SEK flexibel
Aktien Niederlande	Anleihen Global USD-hedged	Mischfonds Sonstige
Aktien Nordeuropa	Anleihen HKD	Mischfonds USD aggressiv
Aktien Norwegen	Anleihen hochverzinslich - Sonsti-ge hedged	Mischfonds USD ausgewogen
Aktien Österreich	Anleihen ILS diversifiziert	Mischfonds USD defensiv
Aktien Osteuropa	Anleihen JPY	Mischfonds USD flexibel
Aktien Osteuropa ohne Russland	Anleihen NOK	Morningstar
Aktien Polen	Anleihen Osteuropa	Rohstoffe - Agrar
Aktien Portugal	Anleihen PLN	Rohstoffe - Diversifiziert
Aktien Russland	Anleihen RMB	Rohstoffe - Edelmetalle
Aktien Schweden Nebenwerte	Anleihen Schwellenländer	Rohstoffe - Sonstige
Aktien Schweden Standardwerte	Anleihen Schwellenländer - lokal	Staatsanleihen EUR

Aktien Schweiz Nebenwerte	Anleihen Schwellenländer EUR-optimiert	Staatsanleihen EUR Kurzläufer
Aktien Schweiz Standardwerte	Anleihen SEK	Staatsanleihen GBP
Aktien Schwellenländer	Anleihen SGD	Staatsanleihen USD
Aktien Singapur	Anleihen Sonstige	Trading - Leveraged/Inverse festverzinslich
Aktien Sonstige islamkonform	Anleihen Sonstige inflationsgesichert	Trading - Leveraged/Inverse Rohstoffe
Aktien Spanien	Anleihen TRY	Trading - Leveraged/Inverse Sonstige
Aktien Südafrika & Namibia	Anleihen USD diversifiziert	Unternehmensanleihen EUR
Aktien Taiwan Standardwerte	Anleihen USD diversifiziert Kurzläufer	Unternehmensanleihen EUR Kurzläufer
Aktien Thailand	Anleihen USD flexibel	Unternehmensanleihen GBP
Aktien Türkei	Anleihen USD hochverzinslich	Unternehmensanleihen USD
Aktien USA flexibel	Anleihen USD inflationsgesichert	Wandelanleihen Asien/Japan
Aktien USA mittelgroß	Branchen: Agrar	Wandelanleihen Europa
Aktien USA Nebenwerte	Branchen: Alternative Energien	Wandelanleihen Global
Aktien USA Standardwerte Blend	Branchen: Andere Sektoren	Wandelanleihen Global CHF-hedged
Aktien USA Standardwerte Growth	Branchen: Biotechnologie	Wandelanleihen Global EUR-hedged
Aktien USA Standardwerte Value	Branchen: Edelmetalle	Wandelanleihen Global GBP-hedged
Aktien USA währungsgesichert	Branchen: Energie	Wandelanleihen Global USD-hedged
Aktien Vietnam	Branchen: Finanzen	Wandelanleihen Sonstige
Aktien weltweit Flex-Cap	Branchen: Gesundheitswesen	Wandelanleihen Sonstige (hedged)
Aktien weltweit islamkonform	Branchen: Immobilienaktien Asien	

Appendix 4: Lipper Fund Categories in Germany[11]

Equity Active Extension Europe	Equity Morocco	Bond EUR Inflation Linked	Mixed Asset CNY Flexible
Equity Active Extension Global	Equity Natural Resources	Bond USD Inflation Linked	Mixed Asset EUR Aggressive - Europe
Equity Active Extension Other	Equity Netherlands	Bond Other Inflation Linked	Mixed Asset EUR Aggressive - EuroZone
Equity Active Extension US	Equity Netherlands Sm&Mid Cap	Bond GBP Inflation Linked	Mixed Asset EUR Aggressive - Global
Equity Argentina	Equity Non-Cyclical Consumer Goods & Services	Bond EUR Corporate	Mixed Asset EUR Balanced - Europe
Equity ASEAN	Equity Nordic	Bond GBP Corporate	Mixed Asset EUR Balanced - EuroZone
Equity Asia Pacific	Equity North America	Bond Global Corporate	Mixed Asset EUR Balanced - Global
Equity Asia Pacific ex Japan	Equity Norway	Bond ILS Corporate	Mixed Asset EUR Conservative - Europe
Equity Asia Pacific Sm&Mid Cap	Equity Nth America Sm&Mid Cap	Bond USD Corporate	Mixed Asset EUR Conservative - EuroZone
Equity Australasia Sm&Mid Cap	Equity Other	Bond ARS Short Term	Mixed Asset EUR Conservative - Global
Equity Australasia	Equity Other	Bond BRL Short Term	Mixed Asset EUR Flexible - Europe
Equity Austria	Equity Pakistan	Bond CAD Short Term	Mixed Asset EUR Flexible - EuroZone
Equity Banks and Other Financials	Equity Pharma & Health Care	Bond CHF Short Term	Mixed Asset EUR Flexible - Global
Equity Basic Industries	Equity Philippines	Bond EUR Long Term	Mixed Asset GBP Aggressive
Equity Belgium	Equity Poland	Bond EUR Medium Term	Mixed Asset GBP Balanced
Equity Biotechnology	Equity Portugal	Bond EUR Short Term	Mixed Asset GBP Conser-

[11] Lipper (2008) *Lipper Global Classification – Extended: Definition Document.* Retrieved from:
http://www.lipperweb.com/docs/support/DataChange/Extended_LGC_Definitions.pdfon 01.11.2012.

			vative
Equity Brazil	Equity Real Estate Europe	Bond Eurozone Long Term	Mixed Asset HKD Aggressive
Equity Canada	Equity Real Estate Global	Bond Eurozone Medium Term	Mixed Asset HKD Balanced
Equity Canada Sm&MidCap	Equity Real Estate Japan	Bond Eurozone Short Term	Mixed Asset HKD Flexible
Equity Chile	Equity Real Estate North America	Bond GBP Short Term	Mixed Asset IDR Balanced
Equity China	Equity Real Estate Other	Bond Global Short Term	Mixed Asset IDR Conservative
Equity Cyclical Consumer Goods & Services	Equity Russia	Bond MAD Short Term	Mixed Asset IDR Flexible
Equity Denmark	Equity Saudia Arabia	Bond MXN Short Term	Mixed Asset IDR Aggressive
Equity Emerging countries of Asia	Equity Singapore	Bond PHP Short Term	Mixed Asset ILS Aggressive
Equity Emerging countries of the European continent	Equity Spain	Bond PHP Medium Term	Mixed Asset ILS Conservative
Equity Emerging countries of the Latin American Continent	Equity Sweden	Bond USD Medium Term	Mixed Asset ILS Flexible
Equity Emerging Markets Other	Equity Sweden Sm&Mid Cap	Bond USD Short Term	Mixed Asset INR Aggressive
Equity Emerging Mkts Asia	Equity Switzerland	Bond Asia Pacific	Mixed Asset INR Balanced
Equity Emerging Mkts Europe	Equity Switzerland Sm&Mid Cap	Bond Europe	Mixed Asset INR Conservative
Equity Emerging Mkts Global	Equity Taiwan	Bond Europe Other	Mixed Asset INR Flexible
Equity Emerging Mkts Latin Am	Equity Taiwan Sm&Mid Cap	Bond Global	Mixed Asset JPY Aggressive
Equity Emerging Mkts Other	Equity Technology, Media and Telecommunication Services	Bond Global AUD Hedged	Mixed Asset JPY Balanced

Equity Eqypt	Equity Telecommunicati-on Services	Bond Global EUR Hedged	Mixed Asset JPY Conser-vative
Equity Europe	Equity Thailand	Bond Global JPY Hedged	Mixed Asset KRW Ag-gressive
Equity Europe ex UK	Equity Turkey	Bond Global USD Hedged	Mixed Asset KRW Balan-ced
Equity Europe Sm&Mid Cap	Equity UAE	Bond Other CHF Hedged	Mixed Asset KRW Con-servative
Equity EuroZone	Equity UK	Bond Other EUR Hedged	Mixed Asset MAD Balan-ced
Equity EuroZone Sm&Mid Cap	Equity UK Sm&Mid Cap	Bond Other USD Hedged	Mixed Asset MAD Con-servative
Equity Finland	Equity Utilities	Bond Other Hedged	Mixed Asset MYR Con-servative
Equity France	Equity Vietnam	Bond CAD High Yield	Mixed Asset MYR Flexible
Equity France Sm&Mid Cap	Bond ARS	Bond EUR High Yield	Mixed Asset MYR Balan-ced – Malaysia
Equity GCC	Bond AUD	Bond Europe High Yield	Mixed Asset MYR Balan-ced – Global
Equity General Industri-als	Bond BRL	Bond GBP High Yield	Mixed Asset NOK Balan-ced
Equity Germany	Bond CAD	Bond Global High Yield	Mixed Asset Other Ag-gressive
Equity Germany Sm&Mid Cap	Bond CHF	Bond USD High Yield	Mixed Asset Other Ba-lanced
Equity Global	Bond CLP	Bond Emerging Markets Global	Mixed Asset Other Con-servative
Equity Global (High UK)	Bond CNY	Bond Emerging Markets Latin Am	Mixed Asset Other Fle-xible
Equity Global ex Japan	Bond DKK	Bond Emerging Markets Other	Mixed Asset PHP Balan-ced
Equity Global ex UK	Bond EUR	Bond Emerging Markets Europe	Mixed Asset PLN Balan-ced
Equity Global ex US	Bond Eurozone	Bond Convertibles Euro-pe	Mixed Asset PLN Conser-vative
Equity Global ex US	Bond GBP	Bond Convertibles Global	Mixed Asset SEK Aggres-

Sm&Mid Cap			sive
Equity Global Income	Bond HKD	Bond Convertibles North America	Mixed Asset SEK Balanced
Equity Global Sm&Mid Cap	Bond IDR	Bond Convertibles Japan	Mixed Asset SEK Conservative
Equity Globally diversified Emerging Markets	Bond ILS	Bond Convertibles Other	Mixed Asset SGD Aggressive
Equity Gold & Precious Metals	Bond INR	Mixed Asset AUD Aggressive	Mixed Asset SGD Balanced
Equity Greater China	Bond INR Government	Mixed Asset AUD Balanced	Mixed Asset SGD Conservative
Equity Hong Kong	Bond JPY	Mixed Asset AUD Conservative	Mixed Asset THB Balanced
Equity India	Bond KRW	Mixed Asset AUD Flexible	Mixed Asset THB Conservative
Equity Indonesia	Bond MAD	Mixed Asset BRL Aggressive	Mixed Asset THB Flexible - Thai
Equity Information Technology	Bond MXN	Mixed Asset BRL Balanced	Mixed Asset THB Flexible - Global
Equity Israel	Bond MYR	Mixed Asset BRL Conservative	Mixed Asset TWD Aggressive
Equity Israel Sm&Mid Cap	Bond NOK	Mixed Asset BRL Flexible	Mixed Asset TWD Balanced
Equity Italy	Bond Other	Mixed Asset CAD Aggressive	Mixed Asset TWD Conservative
Equity Italy Sm&Mid Cap	Bond PHP	Mixed Asset CAD Balanced	Mixed Asset TWD Flexible
Equity Japan	Bond PLN	Mixed Asset CAD Conservative	Mixed Asset USD Aggressive
Equity Japan Sm&Mid Cap	Bond PKR	Mixed Asset CAD Flexible	Mixed Asset USD Balanced - Global
Equity Korea	Bond SEK	Mixed Asset CHF Aggressive	Mixed Asset USD Balanced - North America
Equity Kuwait	Bond SGD	Mixed Asset CHF Balanced	Mixed Asset USD Conservative
Equity Malaysia	Bond THB	Mixed Asset CHF Conser-	Mixed Asset USD Flexible

		vative	- Global
Equity Malaysia Sm&Mid Cap	Bond TND	Mixed Asset CHF Flexible	Mixed Asset USD Flexible - North America
Equity Malaysia/Singapore	Bond TWD	Mixed Asset CLP Conservative	Mixed Asset CNY Flexible
Equity MENA	Bond USD	Mixed Asset CNY Aggressive	Mixed Asset EUR Aggressive - Europe
Equity Mexico	Bond USD Municipal	Mixed Asset CNY Balanced	Mixed Asset EUR Aggressive - EuroZone

Appendix 5: Feri Fund Categories in Germany[12]

Absolute Return Multi Strategy with loss allowance	Convertible Bonds & Options Global EURO-Hedged
Absolute Return Multi Strategy without loss allowance	Dynamic capital preservation funds
Absolute Return Single Strategy without loss allowance	Equity Asia Pacific ex Japan
Balanced Europe balanced	Equity Asia Pacific inc Japan
Balanced Europe conservative	Equity Biotechnology
Balanced Europe dynamic	Equity Central/Eastern Europe
Balanced Global balanced	Equity Emerging Markets
Balanced Global conservative	Equity Euroland
Balanced Global dynamic	Equity Europe
Balanced Global flexible	Equity Europe ex UK
Bond Central-/Eastern Europe	Equity Europe Mid/Small Caps
Bond CHF	Equity Germany
Bond Emerging Markets Local Currencies	Equity Germany Mid/Small Caps
Bond EUR	Equity Global
Bond EUR Corp. High Yield	Equity Health Care Global
Bond EUR Corp. Inv. Grade	Equity India
Bond EUR Global	Equity Japan
Bond EUR Medium-Term	Equity Japan Mid/Small Caps
Bond European Currencies	Equity Latin America
Bond GBP	Equity Middle East and North Africa

---

[12] Feri (2012a) *Feri Fonds Portal.* Retrieved from: http://www.feri-fund-rating.com/Default.aspx?Name=FundsRatingGermany&Content=TopFonds&Lang=en on 31.10.2012.

Bond Global Currencies	Equity North America
Bond Inflation Linked EUR	Equity North America Mid/Small Caps
Bond USD	Equity Sustainability/Ethics World
Bond USD Corp. High Yield	Equity Switzerland
Bonds Emerging Markets Euro-Hedged	Equity Technology Global
Commodities	Equity UK
Convertible Bonds & Options Europe	