Essays on Actively and Passively Managed Financial Products

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum politicarum
(Doktor der Wirtschaftswissenschaft)

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Diplom-Kaufmann Christian Meinhardt

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Dr. Ulrich Kamecke

Gutachter: 1. Prof. Dr. Sigrid Müller
2. Prof. Dr. Lutz Johanning

Tag des Kolloquiums: 23. Juni 2015

Table of Contents

Introduction

Passive investment products and their replication quality

Actively managed mutual funds try to outperform the underlying index by over- or underweighting individual assets. However, the extant literature shows that this strategy is hardly successful. Although actively managed mutual funds often succeed in the short run, they rarely do so in the long run. Consequently, passively managed index products, which attempt to perfectly replicate the underlying index, have seen a robust growth in recent years. Especially exchange-traded funds (ETFs) have attracted enormous cash inflows from investors. At the end of 2000, there were only six ETFs available for investors in Europe as compared to 2,262 ETFs by the end of 2014. Assets under management have increased from EUR 0.65 billion to EUR 553 billion during the same period as reported by BlackRock (2014).

ETFs can be constructed by physical replication, which is a direct investment in the assets that represent the index. In comparison, a synthetically-structured ETF is based on derivatives such as swaps and futures. The synthetic ETF regularly exchanges predefined returns with a counterparty that is often the issuer of the ETF.

The **first study**, "Physical and Synthetic Exchange-Traded Funds: The Good, the Bad, or the Ugly?", investigates this topic by analyzing the replication quality of ETFs and their influencing factors. Theoretically, physical ETFs suffer from higher transaction costs and illiquidity of the replicating index as compared to synthetic ETFs. Consequently, the replication quality of physical ETFs should be worse than that of synthetic ETFs. Additional factors that can influence the replication quality are, for instance, the asset class of the underlying index, the costs measured as both the total expense ratio and the bid-ask spread, the ETFs' liquidity and dividend payments. The three main questions analyzed in this study can be summarized as follows:

- Do synthetic ETFs really replicate their underlying index better than physical ETFs?
- Can we identify differences in the replication quality between equity and fixed-income asset classes?
- Which factors influence the replication quality of ETFs? For instance, this study investigates fund's costs, risk, spread, liquidity and dividends as possible influencing factors.

The results show that physical and synthetic equity ETFs do not differ in replication quality. Thus from an investor's point of view, the tracking errors of both types of ETFs are similar. However, synthetic fixed-income ETFs show smaller tracking errors than physical fixed-income ETFs. Consequently, synthetic ETFs have as good or better tracking errors than physical ETFs. Besides

the construction of ETFs and their underlying asset class, additional factors also significantly influence the ability to replicate the underlying index. Risk, which is defined as the average standard deviation of daily returns, is the factor that universally influences all examined tracking error measures. Costs, dividends and liquidity are some more factors that influence tracking errors of ETFs in varying degrees.

Interestingly, ETFs' successful story cannot be uniformly observed for all passively managed index products. Classical index funds have benefited the least from the enormous growth in recent years and have been almost completely replaced by ETFs. One reason for this is that investors can trade ETFs any time during the stock exchange's trading hours, whereas they can only buy or sell index funds once a day. Kostovetsky (2003), for instance, shows that ETFs have lower management and shareholder transaction fees than index funds. Nonetheless, most studies have not found any difference in the performance between the two index products (e.g. Rompotis (2005), Svetina and Wahal (2008), and Agapova (2011)). Besides index products like ETFs and index funds, index certificates also attempt to perfectly replicate the underlying index, but the extant literature has hardly focused on this derivative till now. Compared to index funds, index certificates have survived notwithstanding the loss of reputation they faced following the collapse of Lehman Brothers.

The **second study**, "Exchange-Traded Funds versus Index Certificates", analyzes this contradiction by comparing the money flows of both ETFs and index certificates within one market. To my knowledge, this is the first study to consider possible dependencies between ETFs and index certificates. Till now, results were only to be found for comparisons between ETFs and index funds (e.g. Svetina and Wahal (2008), and Agapova (2011)). Moreover, this study discusses factors that can have an influence on the relationship between the two index products. These include the ability to replicate the underlying index and a separation of investors into different market niches. The discussed questions can be summarized as follows:

- What are the reasons for the apparent coexistence of both ETFs and index certificates within one market?
- Are ETFs and index certificates substitutes, complements or independent products with regard to their money flow?
- Which index product can best replicate the return of the underlying index?
- Which factors influence the replication quality of ETFs and index certificates? Are there differences between ETFs and index certificates in the context of their influencing factors?

Theoretically, investor demand for two products reacts differently to price changes if the products are substitutes, complements or independent products. If ETFs and index certificates are substitutes, additional inflows to ETFs due to an increased ability to replicate the underlying index, will result in outflows from index certificates or vice versa. However, if ETFs and index certificates are complements, they will both face an additional inflow or outflow although only one of the products will have an increase or decrease in replication quality. Lastly, ETFs and index certificates can be said to be independent products if the money flow of one index product does not influence the money flow of the other index product.

The results show that ETFs and index certificates are complements, though not perfect complements, with regard to their money flow. This result can be explained by similar abilities to replicate the underlying index and a separation of investors into different market niches. Consequently, an increase of the replication quality of ETFs or index certificates results in an increase in inflows to both index products, whereas a decline of the replication quality of one index product increases the flow to the other product. Moreover, the replication quality of both index products significantly changes over time and is primarily influenced by risk and spread.

The evaluation of funds using actively managed investment strategies

Just like ETFs, the universe of mutual funds has increased substantially in the past 20 years. By the end of 2013, the mutual funds' assets under management amounted to EUR 2,100 billion as compared to EUR 158 billion in Germany at the end of 1991 as reported by Bundesverband Investment und Asset Management (BVI). This corresponds to a remarkable growth rate of approximately 60% p.a. Moreover, the number of funds that are available for sale in Germany now is higher than ever before. By the end of 2013, investors had 7,611 different funds to choose from.

Investors are therefore faced with extensive information on numerous funds before decision-making. Information overload makes it difficult for investors to identify the best suitable fund. Thus different questions arise: How can investors recognize whether a fund will outperform the peer group as well as the underlying index in the future? Are there indicators to easily distinguish between funds performing well and poorly?

Fund rating agencies such as Morningstar, Standard & Poor's, Feri Trust and Stiftung Warentest seek to accomplish the requirement of investors to aggregate all available information on funds. They aggregate the information on a fund under one quality label in order to help investors in their decision-making. Literature shows that fund ratings are widely used by investors (e.g. Sirri and Tufano (1998), Del Guercio and Tkac (2008), and Füss et al. (2010)). Investors expect

funds with the best ratings to have high performance. Consequently, these funds obtain enormous cash inflows, whereas lower rated funds frequently suffer from outflows. Hence, the question that arises is whether or not fund rating agencies are really able to provide assistance to investors. Can the future performance of funds be predicted by fund ratings?

Most studies focus on the Morningstar rating and analyze the predictability of the fund rating as well as the impact on the behavior of investors. The results mainly show that the Morningstar rating's ability to predict the future performance of funds is low (e.g. Blake and Morey (2000), Kräussl and Sandelowsky (2007), Antypas et al. (2009), and Füss et al. (2010)). Generally, the performance of top-rated funds does not differ significantly from that of the second-best-rated funds. Only the worst-rated funds frequently have the lowest performance compared to funds with the best ratings.

The **third study**, "Fondsbewertungen in Deutschland – Ein Vergleich der Vorhersagekraft", focuses on this topic by analyzing three different fund ratings in Germany: the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote. To my knowledge, these three fund ratings have hardly been considered in the literature till now. This study discusses the following questions:

- Can these three fund ratings accurately predict the future performance of funds compared to the Morningstar rating?
- If differences in predictability between fund ratings can be identified, which are the factors that can explain these differences? Influencing factors, for instance, could be the valuation process of the fund rating agency, the fund's costs and the behavior of investors.
- Do these three fund ratings influence the behavior of investors?

Fund rating agencies have different methodologies to evaluate funds. Funds can be evaluated quantitatively or qualitatively, albeit combinations between both methods are also possible. Quantitative factors measure the historical performance and the volatility of returns of funds, whereas qualitative factors asses the future performance. Theoretically, fund ratings that combine both quantitative and qualitative factors should be able to better distinguish between well and poorly performing funds compared to fund ratings that only consider quantitative factors (e.g. BVI (2004), BVI (2007a), and Duret et al. (2008)). The Morningstar rating, for instance, relies solely on backward-looking quantitative factors, which might explain the low predictability of the fund rating.

The costs of a fund can also play an important role when analyzing the predictability of fund ratings. Fund returns are influenced differently if the costs differ from each other. Thus significant

differences in performance between fund rating classes before costs can completely disappear after costs (Müller and Weber (2014)). The predictability of fund ratings would then be low.

In addition to the evaluation process of fund ratings and the costs of funds, the behavior of investors could also influence the predictability of fund ratings. If investors primarily invest in the best-rated funds, these funds obtain enormous cash inflows. However, funds cannot go on investing these cash inflows infinitely in assets with the highest performance. Therefore, if funds buy less profitable assets of their benchmark, their high performance today may decrease in the future. Consequently, the predictability of fund ratings would also decline.

Analyses show that the predictability differs significantly between fund ratings and depends, for instance, on the methodology used by the rating agency. The FondsNote, which includes both quantitative and qualitative factors, has the highest predictability, followed by the Feri Trust rating, which is based on quantitative factors. The Finanztest-Bewertung, which also relies on quantitative factors, has the lowest predictability. Moreover, it is shown that the behavior of investors and the costs of the fund can have an impact on predictability. Nevertheless, all three fund rating agencies can only partly distinguish between well and poorly performing funds in future.

The **fourth study**, "Ratings and Performance of German Mutual Funds: A Comparison of Feri Trust, Finanztest, and FondsNote", also analyzes the predictability of the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote but focuses on the combination of these three fund ratings. The investor's decision for a fund not only depends on one single fund rating, but also includes different fund ratings. Consequently, investors mainly invest in funds that have the best fund rating from the different rating agencies. This implies that investors' expectation of fund rating predictability increases when different fund ratings assign funds with the same rating. Thus we can ask:

- Do fund rating combinations increase the predictability to distinguish between well and poorly performing funds in the future?
- Are there differences in predictability when analyzing the future performance of funds and the stability of fund ratings over time?
- Is the forecasting ability of alternative predictors based on backward-looking performance metrics as high as that of the three fund ratings? If this is true, the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote do not deliver added value.

This study uses an extended methodology to evaluate the predictability of the examined fund ratings. Analyses build up on the methodology of Blake and Morey (2000), which is frequently applied in the literature to evaluate the predictability of fund ratings, and use the procedure of Fama and MacBeth (1973). Moreover, this study analyzes whether fund ratings persist over time (e.g. Garnier and Pujol (2007), Duret et al. (2008), and Hereil et al. (2010)).

The results of the predictability of fund ratings are similar in both studies. This speaks for a high validity of the results achieved. Predictability is highest for the Feri Trust rating and the FondsNote compared to the Finanztest-Bewertung. Nonetheless, differences in future performance can only be seen rarely for different performance metrics and post-rating periods. Moreover, this study reveals that predictability can be enhanced by combining the examined fund ratings. However, it still depends on the fund rating combination, the performance metric and the post-rating period. Interestingly, alternative predictors like the Carhart four-factor alpha and the geometric mean of monthly returns are generally not better predictors than the Feri Trust rating, the Finanztest-Bewertung and the FondsNote.

The **fifth study**, "Evaluating the mutual fund rating of Axel Springer and FondsConsult: A case study", investigates the overall predictability of the FondsNote. In this context, it is analyzed whether qualitative factors indeed provide added value in the evaluation process of the fund rating. If qualitative factors are used as part of the evaluation process, they are often similar between fund rating agencies. Only their weights within the final fund rating differ significantly from each other. For instance, FondsNote's qualitative factors are the continuity of the fund management, the consistency of the investment approach and the truth and clarity of the investment product (Axel Springer and FondsConsult (2007)). Feri Trust, Standard & Poor's, Sauren and Scope evaluate funds with similar qualitative factors (BVI (2007b)). This study discusses the following questions:

- Are there differences in predictability between different fund rating categories? Fund rating categories can rely on the asset class (e.g. equity or fixed income) as well as the investment focus (e.g. World, Germany or EUR).
- Do qualitative factors create value added compared to quantitative factors? For instance, we can ask whether the predictability would decrease when the FondsNote would only rely on quantitative factors.
- Do the fund's costs and the investor's behavior decrease the overall predictability of the FondsNote?

The results show that the FondsNote predictability differs between fund rating categories and primarily depends on the performance metric. For instance, best results are found for the Jensen alpha and a three-year post-rating period. However, analyses based on the geometric mean and Sharpe ratio only reveal a low predictability. Qualitative factors, which are used in FondsNote's evaluation process, barely increase the overall predictability of the FondsNote. This result is confirmed by a detailed analysis of both the quantitative and qualitative sub-ratings. Nonetheless, the fund's costs and the behavior of investors on fund ratings influence the predictability of the FondsNote.

References

Agapova, A. (2011). Conventional mutual index funds versus exchange-traded funds. Journal of Financial Markets 14 (2), pp. 323-343.

Antypas, A., Caporale, G. M., Kourogenis, N., Pittas, N. (2009). Selectivity, market timing and the Morningstar star-rating system. Working paper, No. 874, DIW Berlin, April.

Axel Springer, FondsConsult (2007). FondsNote. Das Fonds-Rating von Axel Springer Finanzen Verlag und FondsConsult. Axel Springer Finanzen Verlag & FondsConsult Research AG, Munich.

BlackRock (2014). ETP landscape. Industry highlights. BlackRock Inc., December 2014.

Blake, C. R., Morey, M. R. (2000). Morningstar ratings and mutual fund performance. The Journal of Financial and Quantitative Analysis 35 (3), pp. 451-483.

BVI (2004). Transparenz-Standards für die Beurteilung von Investmentfonds durch Ranking-/Ratingagenturen. Ranking-/Rating-Transparenz-Standards, RRTS, Bundesverband Investment und Asset Management e.V, December.

BVI (2007a). Was sagen Rankings und Ratings aus? BVI Jahrbuch 2007, Bundesverband Investment und Asset Management e.V., pp. 65-71.

BVI (2007b). Ranking-/Rating-Transparenz-Standard – Antwortsammlung. Bundesverband Investment und Asset Management e.V., 3rd edition, October 2007.

Del Guercio, D., Tkac, P. A. (2008). Star power: the effect of Morningstar ratings on mutual fund flow. The Journal of Financial and Quantitative Analysis 43 (4), pp. 907-936.

Duret, A., Hereil, P., Mitaine, P., Moussavi, N., Roncalli, T. (2008). Fund rating systems and performance predictability. Working paper, University of Evry, April.

Fama, E. F., MacBeth, J. D. (1973). Risk, return, and equilibrium: empirical tests. The Journal of Political Economy 81 (3), pp. 607-636.

Füss, R., Hille, J., Rindler, P., Schmidt, J., Schmidt, M. (2010). From rising stars and falling angels: on the relationship between the performance and ratings of German mutual funds. Journal of Wealth Management 13 (1), pp. 75-90.

Garnier, O., Pujol, T. (2007). Can today's stars be used to read the stars of the future? Working paper, No. 3, March.

Hereil, P., Mitaine, P., Moussavi, N., Roncalli, T. (2010). Mutual fund ratings & performance persistence. White paper 3, Quant Research by Lyxor, June.

Kostovetsky, L. (2003). Index mutual funds and exchange-traded funds. The Journal of Portfolio Management 29 (4), pp. 80-92.

Kräussl, R., Sandelowsky, R. (2007). The predictive performance of Morningstar's mutual fund ratings. Working paper, VU University Amsterdam, August.

Müller, S., Weber, M. (2014). Evaluating the rating of Stiftung Warentest: how good are mutual fund ratings and can they be improved? European Financial Management 20 (2), pp. 207-235.

Rompotis, G. G. (2005). An empirical comparing investigation on exchange traded funds and index funds performance. Working paper, December.

Sirri, E. R., Tufano, P. (1998). Costly search and mutual fund flows. The Journal of Finance 53 (5), pp. 1589-1622.

Svetina, M., Wahal, S. (2008). Exchange traded funds: performance and competition. Working paper, November.

Part I:
Physical and Synthetic Exchange-Traded Funds: The Good, the Bad, or the Ugly?

Christian Meinhardt, Sigrid Mueller, and Stefan Schoene

In: Journal of Investing 24 (2), pp. 35-44, Summer 2015.

Abstract

This paper focuses on the replication process of exchange-traded funds (ETFs). It compares the tracking ability of ETFs based on physical replication of their benchmark indices to those of synthetic ETFs. Synthetic ETFs rely on derivatives such as swaps. For ETFs listed at the Frankfurt Stock Exchange we show that both categories of ETFs suffer from high tracking errors. Contrary to conventional wisdom, synthetic equity ETFs are not different in terms of tracking errors from their physical counterparts. However, synthetic fixed-income ETFs have lower tracking errors than physical fixed-income ETFs. Thus synthetic ETFs have as good or better tracking errors than physical ETFs. We identify different factors influencing tracking errors.

JEL Classification: G11, G12, G19, G23

Keywords: Exchange-traded fund, physical replication, synthetic ETF, tracking error, systemic risk, serial correlation

Part II:
Exchange-Traded Funds versus Index Certificates

Abstract

This study examines the coexistence of two index products, namely ETFs and index certificates, within one market by analyzing the relationship between each other's money flow. Evidence shows that ETFs and index certificates are complements, but not perfect complements to each other, which means that inflows to both index products are correlated positively. This effect can be explained by similar tracking abilities and a separation of investors into different market niches. Nonetheless, this study also finds strong evidence that tracking abilities significantly change over time. These changes can be explained by risk and spread.

JEL Classification: C33, G11, G12, G19, G23

Keywords: ETF, index certificate, tracking error, tracking difference, influencing factor, panel-regression

I would like to thank David Bosch, Sigrid Müller, Elina Pradkhan, and Stefan Schöne from the Institute of Finance for valuable comments and suggestions. Moreover, I am grateful to Lutz Johanning and the European Derivatives Group (EDG) for providing me with data on open interest of index certificates.

1 Introduction

There are basically two types of investment strategies for investing in a basket of securities. An actively managed strategy tries to outperform an underlying index, while a passively managed strategy aims for an exact replication of the index. In recent years, products using passive investment strategies have seen a robust growth. Especially in the American and German markets, exchange-traded funds (ETFs) benefit from this development by attracting enormous cash inflows from investors. By the end of 2012, the market volume of ETFs amounted to USD 1,933 billion worldwide, which included USD 1,349 billion in the U.S. and USD 367 billion in Europe as reported by BlackRock (2012).

In addition to ETFs, other index products such as index funds and index certificates seek to perfectly replicate the underlying index. However, these index products have benefited to a lesser extent from the popularity of passive investment strategies in the past years. Especially, index funds have even gradually been replaced by ETFs. One might assume that ETFs have also started to replace index certificates, especially after their loss of reputation during the collapse of Lehman Brothers. On the contrary, index certificates still retain their popularity, especially in Germany.

This study deals with this apparent contradiction by analyzing the flows to ETFs and index certificates and their influence on each other's flow. To my knowledge, this is the first ever study to investigate this topic. Detailed analyses are especially suited for the German market because investors can easily choose between ETFs and index certificates.

Analyses also focus on factors that determine the relationship of flows between ETFs and index certificates. For investors, it is important to know which financial product is best suited to replicate a preferred index because investment decisions are mainly based on the tracking ability. This study deals with this topic by analyzing changes in the tracking ability of ETFs and index certificates replicating 23 major European indices over time. Moreover, this study analyzes the influence of different factors on the tracking ability.

Both ETFs and index certificates have the same objective, namely replicating the underlying index as closely as possible. Thus, investors can easily participate in the return of an index and diversify their portfolio risks at low cost. Moreover, since ETFs and index certificates can be traded all day long, their replication quality is easily comparable using index certificates with unlimited maturity and no capital guarantee. The main difference between ETFs and index certificates certainly depends on their construction.

ETFs can be constructed by directly investing in the assets that represent the index, and this is called physical replication. Compared to that, a synthetically structured ETF is based on derivatives such as swaps and futures. The synthetic ETF regularly exchanges predefined returns with a counterparty that is often the issuer of the ETF. Similar to synthetic ETFs, physical ETFs face a counterparty risk since they generate an additional income due to lending of securities. However, neither synthetic nor physical ETFs bear any default risk of the issuer, because the assets of the fund are held in a separate account. Compared to ETFs, the redemption of index certificates depends on the solvency of the issuer, because this type of index product is constructed as a bearer bond. For further differences and commonalities between the two index products see Klein and Kundisch (2008) and Johanning et al. (2011).

Key results of this study can be summarized as follows: Evidence shows that ETFs and index certificates are complements, but they are not perfect complements to each other. This effect can be explained by similar tracking abilities and a separation of investors into different market niches. There is also strong evidence that tracking abilities change significantly over time. Moreover, this study finds that these changes can be explained by an increase in the impact as well as an increase in the value of risk and spread during the examination period. Nonetheless, both ETFs and index certificates suffer from high tracking errors, but not from high performance differences.

This paper is structured in the following way: Section 2 briefly summarizes the existing literature on ETFs and index certificates. Section 3 describes the data set used in this study and the methodology applied to derive the impact of flows on each other's flow, the replication quality and their influencing factors. Section 4 presents the empirical results and Section 5 concludes with summarizing the findings.

2 Literature Review

Contrary to the literature on ETFs, studies on index certificates are rather scarce. Many studies concentrate on the comparison of the performance and tracking quality between ETFs and index funds. Elton et al. (2002) are the first to note an underperformance of the ETF Spiders (SPDR) in comparison to the S&P 500 index and an index fund. The principal causes of the underperformance of the ETF Spiders are the management fee and the loss on dividend reinvestment. Similar results are provided by Blitz et al. (2012) and Gastineau (2004). Blitz et al. (2012) also describe dividend withholding taxes as an important determinant of underperformance. Agapova (2011), Rompotis (2005) and Svetina and Wahal (2008) find that the performance of ETFs is indistinguishable from directly competing index funds. Poterba and Shoven (2002) compare the pre- and post-

tax returns of the SPDR with the Vanguard 500 index fund. The results imply that both the ETF and the index fund have similar performances. Agapova (2011) also notices that ETFs, on the average, have smaller tracking errors as compared to index funds, whereas Rompotis (2005) does not find any difference in tracking errors.

Graf (2001) is the first to compare the tracking ability of the first ETF listed at the Swiss Stock Exchange with several index funds and index certificates. The results suggest that tracking errors of the ETF and the index funds are lower than that of index certificates. Return differences between index products and the underlying index are similar but negative. Only the ETF outperforms the index. Graf (2001) mentions the liquidity of the underlying index as a key factor for tracking ability.

Schmidhammer et al. (2011) focus on intraday analyses of ETFs, index certificates and corresponding DAX 30 index prices and DAX 30 index futures prices. The performance of the analyzed index products does not differ from the underlying index. Indeed, Schmidhammer et al. (2011) determine that the replication quality of ETFs is higher than that of index certificates in the short run. Moreover, futures contracts contribute a significant proportion to contemporaneous price quotes of ETFs and index certificates.

Klein and Kundisch (2009b) also focus on intraday trades with respect to the price-setting behavior of index certificates replicating the DAX 30 index. They conclude that the differences in return between index certificates and the index can be explained by the spread. The results indicate that prices of open-end index certificates do not follow a life-cycle. Based on a decision-theoretical model, Klein and Kundisch (2008) analyze ETFs and index certificates in Germany depending on returns, risks and tax effects. Moreover, Klein and Kundisch (2009a) find a decrease in tracking errors of the first ETF replicating the DAX 30 index and ten index certificates from 2001 to 2006.

Besides the tracking ability of ETFs and index certificates, only some studies concentrate on the aspects of competition and their influence on each other's flow. Svetina and Wahal (2008) point out that ETFs generally do not directly compete with index funds on a common index. Launching new ETFs on an index permanently reduces net flows of existing index products and improves replication quality. Meinhardt et al. (2015) compare the replication quality of synthetically and physically replicating ETFs. In this context, it is even shown that the two groups only rarely compete on a common index. Agapova (2011) compares fund flows of ETFs and index funds, and concludes that the index products are substitutes, but they are not perfect substitutes. A clientele effect explains the coexistence of ETFs and index funds by dividing the two index products into different market niches.

This study contributes to the existing literature by empirically analyzing the tracking ability of competing ETFs and index certificates and their influencing factors for one entire time period as well as rolling time periods. Compared to the literature on index certificates, the present sample consists of a sufficiently large data set and examination period to obtain plausible results. Furthermore, it is the first study, to my knowledge, that statistically examines whether ETFs and index certificates that track a common index are complements, substitutes or independent products.

3 Data and Methodology

3.1 Data

This study compares competing ETFs and index certificates replicating 23 different European indices. Analyses of the flows to ETFs and index certificates as well as their influence on each other's flow are based on 91 ETFs and 100 index certificates from 01/01/2007 to 12/31/2012. Tracking abilities and their influencing factors are analyzed for 72 ETFs and 113 index certificates from 06/01/2009 to 12/31/2012.

The list of ETFs is from BlackRock, whereas the list of index certificates is primarily collected from Scoach, the European stock exchange for structured products in Frankfurt. Delisted index certificates are gathered from the European Derivatives Group (EDG). This study only takes into account accumulating index certificates that attempt to identically replicate the return of their underlying index. Inverse, leveraged, distributing as well as bonus and discount structured products are not included. The list of ETFs includes both accumulating and distributing ETFs. Dividends of distributing ETFs are reinvested in the ETF at the day of distribution, which enables comparisons between ETFs and index certificates.

EDG provided me with data on monthly open interest of index certificates as well as the market volume of the certificate universe in Germany as of 01/01/2007. The EDG data set covered 70% of all index certificates in Germany at the beginning of the examination period and soon increased to 95%. This study analyzes monthly flows to ETFs and index certificates as of 01/01/2007 to obtain the most plausible results. The total net assets of ETFs are collected from ETF providers and Bloomberg.

Daily bid and ask quotes of index certificates are available from Scoach as of 06/01/2009. The quotes are manually collected at the time of the last index pricing in order to eliminate the time gap between the prices of the index certificates and indices. The daily last Xetra prices of ETFs at the Frankfurt Stock Exchange and the last index prices (net and gross return) are downloaded

from the homepages of index providers and Bloomberg. The total expense ratios, the volume of securities lending and dividend payments are obtained from BlackRock and ETF providers.

Table 1 presents descriptive statistics of the number of analyzed index products, means of monthly assets under management (AUM) as well as means of monthly changes in flows, separated into ETFs and index certificates and grouped by the index. The majority of analyzed ETFs replicate the EURO STOXX 50, whereas most index certificates replicate the DAX 30. In terms of AUM, ETFs are on the average larger than index certificates for all 23 indices. Furthermore, flows to ETFs are on the average positive and significantly higher than flows to index certificates. Index certificates even attain negative flows on an average.

Table 1: Descriptive Statistics

	ETFs				Certificates			
Index	$\#_{Flow}$	$\#_{TE}$	**AUM**	**Flow**	$\#_{Flow}$	$\#_{TE}$	**AUM**	**Flow**
DAX 30	7	6	1465.86	30.6138	32	32	63.62	-0.6234
EURO STOXX 50	12	10	958.66	5.5066	5	5	12.49	-0.1975
EURO STOXX Select Dividend 30	3	3	226.39	1.2827	4	4	5.75	-0.0385
STOXX Europe 50	3	2	326.48	-1.5602	3	2	16.02	-0.2340
STOXX Europe 600	4	4	118.52	2.2793	3	2	19.11	0.5040
STOXX Europe 600 Automobiles & Parts	3	2	21.35	-0.0900	4	4	0.03	0.0000
STOXX Europe 600 Banks	4	3	138.67	3.7600	5	5	0.16	-0.0045
STOXX Europe 600 Basic Resources	4	4	88.35	0.3743	4	4	0.04	0.0014
STOXX Europe 600 Chemicals	3	2	35.88	-0.1349	2	4	0.33	-0.0146
STOXX Europe 600 Construction & Materials	3	2	23.27	-0.0749	2	4	0.03	0.0001
STOXX Europe 600 Financial Services	3	2	19.02	-0.0092	2	4	0.38	-0.0172
STOXX Europe 600 Food & Beverage	3	3	48.98	0.5774	4	4	0.04	-0.0004
STOXX Europe 600 Health Care	4	3	92.10	-0.1359	4	4	0.06	-0.0007
STOXX Europe 600 Industrial Goods & Services	4	3	32.32	0.2992	3	3	0.02	-0.0015
STOXX Europe 600 Insurance	4	3	50.47	-0.0085	3	4	0.10	-0.0043
STOXX Europe 600 Media	3	2	24.61	-0.1453	1	4	0.02	-0.0004
STOXX Europe 600 Oil & Gas	3	3	98.18	0.4380	4	4	0.05	-0.0005
STOXX Europe 600 Personal & Household Goods	3	2	34.42	0.1268	2	3	0.01	0.0000
STOXX Europe 600 Retail	3	2	29.63	-0.3313	3	3	0.03	0.0008
STOXX Europe 600 Technology	4	3	35.50	-0.1805	2	3	0.01	-0.0009
STOXX Europe 600 Telecommunications	4	3	67.38	0.0448	2	4	0.11	0.0024
STOXX Europe 600 Travel & Leisure	3	2	14.08	-0.0492	2	3	0.01	0.0000
STOXX Europe 600 Utilities	4	3	61.82	-0.0775	4	4	0.10	0.0023
All indices	91	72	299.04	3.3614	100	113	22.31	-0.2037

The table presents descriptive statistics of ETFs and index certificates. The overall universe of ETFs and index certificates is divided into two subsamples. $\#_{Flow}$ and $\#_{TE}$ are the numbers of used ETFs and index certificates for each index, respectively. Analysis of fund flows is based on $\#_{Flow}$ and tracking errors, tracking differences and influencing factors are based on $\#_{TE}$. Means are reported for assets under management (AUM) in millions of euro, which is the total net assets for ETFs and the open interest multiplied with the price for index certificates, and flow which is the net flow to an ETF or index certificate in millions of euro. Both AUM and flow are based on $\#_{Flow}$ and monthly data.

3.2 Methodology

Substitutes, complements or independent products

Economic theory implies that investor demand for two products is influenced differently by price movements if the two products are substitutes, complements or independent products. For ETFs and index certificates, the demand can be measured by the flows to ETFs and index certificates, and is primarily determined by the ability of the products to track the underlying index.

If ETFs and index certificates are substitutes, they will influence each other's flow negatively. For example, additional inflows to ETFs due to a better tracking ability will result in outflows from index certificates or vice versa. Moreover, perfect substitution takes place if increasing flows to index certificates negatively affect ETF inflows and positively affect outflows from ETFs. However, if ETFs and index certificates are complements, the flows to both index products are positively correlated. For example, if only one product enhances their tracking ability, both index products will simultaneously gain from an increase in investor demand. ETFs and index certificates are perfect complements if flows to index certificates positively influence ETF inflows and negatively influence ETF outflows.

As information on the flows of individual investors is not available, the investor demand for ETFs and index certificates is determined by aggregated flows. I use the methodology of Sirri and Tufano (1998) to calculate aggregated flows to ETFs and index certificates, respectively.

$$Flow_{i,t} = AUM_{i,t} - AUM_{i,t-1} \cdot (1 + R_{i,t}) \tag{1}$$

The monthly aggregated flow ($AUM_{i,t}$) of ETF and index certificate i at time t is measured by the assets under management and the open interest. $R_{i,t}$ is the return of the ETF and index certificate over the previous month. In this way, flows to ETFs and index certificates are corrected for changes in flows due to the returns from $t-1$ to t.

Based on the study of Agapova (2011), the influence on each other's flow is tested with a system of equations (2) and (3). I use the seemingly unrelated regressions procedure (SUR) because the variables flows to ETFs (*FlowETF*) and flows to index certificates (*FlowCert*) are included in both equations as dependent and explanatory variables. Fixed effects are controlled by year and index dummies. For the sake of completeness, the system is also tested with an ordinary least squares regression including year and index dummies.

$$\begin{aligned} FlowETF_{i,t} = \alpha_{i,t} &+ \beta_1 FlowCert_{i,t} + \beta_2 FlowCert_{i,t-1} + \beta_3 FlowCert_{i,t-2} \\ &+ \beta_4 FlowETF_{i,t-1} + \beta_5 FlowETF_{i,t-2} + \beta_6 Ret_{i,t} + \beta_7 Ret_{i,t-1} \\ &+ \beta_8 Ret_{i,t-2} + \beta_9 Costs_{i,t} + \beta_{10} LogMarketCap\_ETF_{i,t} \end{aligned}$$

$$+\beta_{11} FlowUniverse\_ETF_{i,t} + \varepsilon_{i,t} \quad (2)$$

$$\begin{aligned} FlowCert_{i,t} = \alpha_{i,t} &+ \beta_1 FlowETF_{i,t} + \beta_2 FlowETF_{i,t-1} + \beta_3 FlowETF_{i,t-2} \\ &+\beta_4 FlowCert_{i,t-1} + \beta_5 FlowCert_{i,t-2} + \beta_6 Ret_{i,t} + \beta_7 Ret_{i,t-1} \\ &+\beta_8 Ret_{i,t-2} + \beta_9 LogMarketCap\_Cert_{i,t} \\ &+\beta_{10} FlowUniverse\_Cert_{i,t} + \varepsilon_{i,t} \end{aligned} \quad (3)$$

Explanatory variables include lagged flows to ETFs and index certificates, current and lagged returns and the logarithm of the market capitalization of all ETFs and index certificates replicating the same index. Both the returns and the logarithm of market capitalization are calculated as the equally-weighted average of all index products on a common index, respectively. The equally-weighted average of current total costs per share and the flow of the market volume of all ETFs in Germany are used in equation (2). The flow of the market volume of all index certificates in Germany is used in equation (3).

The results of the regressions can be interpreted in the following way. If both β_1 coefficients are significantly positive, ETFs and index certificates are complements. If one β_1 coefficient is negative, ETFs and index certificates are substitutes implying, for example, that an increase of the assets under management of ETFs lead to a decrease of the assets under management of index certificates (Yan (2006) and Agapova (2011)).

Current and lagged returns are performance measures that show the desirability of ETFs and index certificates. Total costs of ETFs measure the expenses of investors and should have a negative relation to ETF flows. A possible size effect is controlled by the market capitalization of ETFs and index certificates. Flows to the universe of ETFs and index certificates measure the overall attractiveness of the index products and should be positively related to flows to ETFs and index certificates.

The SUR approach is also tested with ETF flows (*FlowETF*) divided into ETF inflows (*FlowETF_in*) and ETF outflows (*FlowETF_out*) to analyze whether substitution or complementarity is only based on inflows or on outflows. I use two systems of equations with *FlowETF_in*, *FlowETF_out* and *FlowCert* as dependent variables. Each equation is controlled by year and index dummies.

Tracking error and tracking difference

Investor demand can be mainly determined by the ability to track the underlying index. A substitution effect can occur if tracking abilities of ETFs and index certificates differ from each other. Thus, both index products would have the incentive to enhance the tracking ability by reducing costs or enhancing returns compared to index returns. The index product with an advantage in

tracking ability will attract more investor inflow, whereas the other will have additional outflow. If tracking abilities of the respective ETFs and index certificates are similar, other factors like the default risk of index certificates can explain a substitution effect.

Complementarity between both index products can be explained by ETFs and index certificates that have similar tracking abilities. Moreover, the tracking ability of index certificates can even be higher than that of ETFs, compensating for the default risk of index certificates. Thus, the incentive of ETFs and index certificates to enhance the tracking ability will be smaller when comparing these two index products only.

The tracking ability can be measured by the tracking error and the differences in performances between the index product and the index. In this study, I analyze ETFs and index certificates by the standard deviation of return differences, a method that is predominantly used for measuring tracking errors in literature and practice. When comparing performances, I rely on the average difference of returns between the index product and the underlying index. Johnson et al. (2013) introduce the term tracking difference for the difference in returns. I use this definition for the sake of simplicity. Means of tracking errors and tracking differences are calculated for the entire time period and for daily rolling one-year periods, respectively. Separating the whole examination period into daily rolling one-year periods results in 660 sub-periods starting from 06/01/2009 till 12/31/2011.

Pope and Yadav (1994) find that the use of daily returns might result in an overestimation of tracking errors. Thus, return differences between index product and the index are corrected for serial correlation using the Cochrane-Orcutt procedure. In some studies, weekly data is used to overcome the problem of correlated return differences. However, this practice is not suitable in the context of this study because the longest time period for adjustments is three weeks for a single ETF and index certificate. Moreover, tracking errors and tracking differences of distributing ETFs might be higher than for accumulating ETFs and index certificates because the return difference between index product and index increases at the moment of distribution. For this reason dividends are reinvested immediately which is necessary for 24 ETFs paying out dividends regularly.

Factors influencing tracking errors

Several factors can influence the tracking ability. Frino and Gallagher (2001) mention dividends, changes in index composition and net cash flows as driving factors. Kostovetsky (2003) points out that expenses like management fees and transaction costs can negatively influence tracking ability. Moreover, Milonas and Rompotis (2006) analyze risk as the standard deviation of

daily returns, whereas Shin and Soydemir (2010) use changes in exchange rate. Both risk and changes in exchange rate negatively affect tracking errors.

In this study, I analyze the influence of six different factors on tracking errors. These are for both ETFs and index certificates: risk, spread and volume. Additional analyzed factors for ETFs are dividends, securities lending and total costs. Influencing factors on tracking differences are not considered separately because factors influencing tracking errors also influence tracking differences.

Previous studies have analyzed the influence of the mean of each factor on tracking error using cross-sectional regressions. This way, differences in factors are lost over time. For this reason, I analyze the effect on the tracking errors of ETFs and index certificates by using both a pooled-cross-sectional regression as well as a panel-regression with fixed-effects.

$$\begin{aligned} TE\_ETF_{i,t} = \alpha_{i,t} &+ \beta_1 Risk_{i,t} + \beta_2 Spread_{i,t} + \beta_3 Volume_{i,t} + \beta_4 Dividends_{i,t} \\ &+ \beta_5 Costs_{i,t} + \beta_6 SecLending_{i,t} + \beta_7 Risk_{i,t} \cdot D_t + \beta_8 Spread_{i,t} \cdot D_t \\ &+ \beta_9 Volume_{i,t} \cdot D_t + \beta_{10} Dividends_{i,t} \cdot D_t + \beta_{11} Costs_{i,t} \cdot D_t \\ &+ \beta_{12} SecLending_{i,t} \cdot D_t + \varepsilon_{i,t} \end{aligned} \tag{4}$$

$$\begin{aligned} TE\_Cert_{i,t} = \alpha_{i,t} &+ \beta_1 Risk_{i,t} + \beta_2 Spread_{i,t} + \beta_3 Volume_{i,t} + \beta_4 Risk_{i,t} \cdot D_t \\ &+ \beta_5 Spread_{i,t} \cdot D_t + \beta_6 Volume_{i,t} \cdot D_t + \varepsilon_{i,t} \end{aligned} \tag{5}$$

Tracking errors are calculated as one-year standard deviation of return differences for each ETF and certificate i at time t. Besides, the analysis is repeated with the mean of tracking errors based on index i at time t. Fixed effects are controlled by including dummies based on time and index products (ETFs and index certificates) or indices.

Explanatory variables are: risk, which is measured as the standard deviation of daily returns; spread, which is measured as the average of daily relative spreads; volume, measured as the natural logarithm of the average daily trading volume; dividends, measured as the ratio of average dividend and average trading price; securities lending, measured as the ratio of net 12 month securities lending revenue and average assets under management over the same period; total costs, which is measured as the average daily total costs per share. All analyzed factors are assumed to influence the tracking ability negatively except for the factor of volume that affects the tracking ability positively. Moreover, D_t represents a dummy variable which separates the examination period into two sub-periods to control the differences in influence of analyzed factors after July 2011 when the financial crisis spilled over to Italy and Spain.

4 Results

4.1 Funds Flow Analysis

Table 2, Panel A, shows the influence of flows to ETFs and index certificates on each other's flow. The results suggest that β_1 coefficients are significantly positive in both equations with the SUR approach indicating that ETFs and index certificates are complements. Only the general pooled OLS regression yields insignificant β_1 coefficients.

Moreover, the current flows are influenced by lagged flows. An increase of lagged flows to ETFs decreases the current ETF flows but does not increase the current flows to index certificates within all the testing parameters. In comparison, lagged flows to index certificates positively influence the current flows to index certificates and negatively influence the current ETF flows. Current ETF flows are also positively related to the size of the ETF and to an increase in flows to the ETF universe. Interestingly, total costs have no immediate impact on ETF flows. The current flows to index certificates are negatively related to their size in the SUR fixed-year approach. All the other regressions yield insignificant results. Finally, the flow of the entire universe of certificates is statistically significant at the 10% level with all the test specifications.

Is the determined complementarity between ETFs and index certificates a result of inflows or outflows? Are ETFs and index certificates even perfect complements, which influence both inflows and outflows? Table 2, Panel B, reports the results of ETF inflows and outflows and the net flows of index certificates.

Flows to index certificates significantly influence ETF inflows and outflows in a positive manner. This is consistent with the predictions for inflows, but inconsistent with the forecast for outflows. Nevertheless, ETF inflows are positively related to the flows to index certificates. The influence of ETF outflows on flows to the index certificates is certainly positive with the SUR fixed-year approach and therefore inconsistent with the earlier predictions. The results confirm the complementarity of ETFs and index certificates, but do not confirm their perfect complementarity. Consequently, an increase of the tracking ability of ETFs or index certificates can result in an increase of inflows to both index products. On the contrary, a decrease in the tracking ability of one index product can increase the flows to the other product.

Independent of the tracking ability of the index products, the complementarity of inflows can be explained by a clientele effect. Investors of ETFs and index certificates invest at the same time but they might have different preferences with regard to the duration for which an investment is held before it is liquidated. ETFs are mainly used by investors as a buy-and-hold strategy to increase the assets under management permanently (e.g. Agapova (2011)). Compared to ETF inves-

tors, Johanning et al. (2011) point out that investors of index certificates mainly invest for a short period of time, wherefore the flows to index certificates fluctuate more strongly. This clientele effect separates investors into different market niches and can additionally explain the complementarity between the inflows to ETFs and index certificates.

Table 2: Substitutes, Complements or Independent Products

	FlowETF			FlowCert		
Panel A	**OLS**	**$SUR_{Year/Index}$**	**SUR_{Year}**	**OLS**	**$SUR_{Year/Index}$**	**SUR_{Year}**
Intercept	**-970.276*** **	**-982.402*** **	**-221.463*** **	16.353	18.910	**7.666** **
	(-3.129)	(-3.227)	(-3.192)	(0.946)	(1.114)	(2.420)
FlowCert	0.888	**1.684*** **	**1.147** **			
	(1.632)	(3.153)	(2.140)			
$FlowCert_{t-1}$	**-0.871***	**-0.912***	**-1.169** **	**0.055***	**0.057** **	**0.071** **
	(-1.683)	(-1.796)	(-2.293)	(1.885)	(1.986)	(2.475)
$FlowCert_{t-2}$	-0.464	-0.645	-0.768	**0.229*** **	**0.229*** **	**0.242*** **
	(-0.881)	(-1.248)	(-1.474)	(7.885)	(8.049)	(8.465)
FlowETF				0.003	**0.005*** **	**0.003** **
				(1.537)	(3.055)	(2.006)
$FlowETF_{t-1}$	**-0.213*** **	**-0.213*** **	**-0.195*** **	0.000	0.001	0.000
	(-6.968)	(-7.085)	(-6.461)	(0.102)	(0.425)	(-0.284)
$FlowETF_{t-2}$	**-0.108*** **	**-0.108*** **	**-0.090*** **	0.000	0.000	-0.001
	(-3.542)	(-3.600)	(-3.005)	(-0.024)	(0.141)	(-0.475)
Return	-122.824	-120.839	-116.588	-2.594	-2.440	-2.924
	(-1.288)	(-1.290)	(-1.240)	(-0.493)	(-0.472)	(-0.565)
$Return_{t-1}$	-30.270	-31.214	-9.229	1.851	1.937	1.480
	(-0.326)	(-0.342)	(-0.101)	(0.352)	(0.375)	(0.286)
$Return_{t-2}$	10.023	5.282	24.149	5.547	5.541	5.324
	(0.108)	(0.058)	(0.265)	(1.061)	(1.079)	(1.034)
Costs	532.653	529.667	-33.802			
	(0.774)	(0.785)	(-0.117)			
LogMarketCap_ETF	**242.126*** **	**245.137*** **	**87.059*** **			
	(3.187)	(3.286)	(4.936)			
LogMarketCap_Cert				-4.016	-4.701	**-3.252*** **
				(-0.935)	(-1.114)	(-3.250)
FlowUniverse_ETF	**0.000*** **	**0.000*** **	**0.000*** **			
	(2.820)	(2.868)	(3.128)			
FlowUniverse_Cert				**0.359***	**0.358***	**0.353***
				(1.893)	(1.928)	(1.886)
R^2	0.10	0.09	0.07	0.11	0.11	0.10
Adj. R^2	0.06	0.06	0.06	0.08	0.08	0.09
Year dummies	yes	yes	yes	yes	yes	yes
Index dummies	yes	yes	no	yes	yes	no

Panel B	**FlowETF_in**		**FlowETF_out**		**FlowCert**	
Intercept	**-569.772****	**-306.111*****	-104.897	**89.100*****	17.700	**7.350****
	(-2.438)	(-5.664)	(-0.770)	(2.872)	(1.042)	(2.239)
FlowCert	**1.622*****	**0.862****	**0.417***	**0.600****		
	(3.973)	(2.060)	(1.737)	(2.499)		
FlowETF_in					**0.009*****	**0.004***
					(3.870)	(1.764)
FlowETF_out					0.004	**0.008****
					(1.050)	(2.200)
FlowETF_in$_{t-1}$	0.023	0.038			-0.002	-0.002
	(0.797)	(1.298)			(-0.901)	(-1.102)
FlowETF_in$_{t-2}$	0.022	0.041			-0.002	-0.003
	(0.824)	(1.473)			(-1.164)	(-1.474)
FlowETF_out$_{t-1}$			**-0.508*****	**-0.523*****	0.004	0.004
			(-22.231)	(-22.762)	(1.049)	(1.212)
FlowETF_out$_{t-2}$			**-0.172*****	**-0.183*****	0.005	0.004
			(-6.912)	(-7.269)	(1.464)	(1.324)
FlowCert$_{t-1}$	-0.435	**-0.777***	**-0.679*****	**-0.623*****	**0.062****	**0.079*****
	(-1.118)	(-1.953)	(-2.976)	(-2.720)	(2.161)	(2.722)
FlowCert$_{t-2}$	**-1.462*****	**-1.639*****	0.040	0.030	**0.245*****	**0.250*****
	(-3.688)	(-4.023)	(0.171)	(0.128)	(8.440)	(8.576)
Return	-90.325	-82.586	-28.328	-30.019	-2.110	-2.877
	(-1.261)	(-1.126)	(-0.674)	(-0.713)	(-0.409)	(-0.557)
Return$_{t-1}$	-25.848	-8.828	3.865	2.584	1.848	1.361
	(-0.370)	(-0.124)	(0.094)	(0.063)	(0.358)	(0.263)
Return$_{t-2}$	11.000	27.615	3.292	1.405	5.303	5.196
	(0.158)	(0.388)	(0.081)	(0.034)	(1.035)	(1.010)
Costs	381.624	-244.974	-61.286	170.493		
	(0.740)	(-1.084)	(-0.202)	(1.314)		
LogMarketCap_ETF	**181.877*****	**149.981*****	-9.468	**-61.333*****		
	(3.177)	(10.986)	(-0.284)	(-7.806)		
LogMarketCap_Cert					-4.634	**-3.069*****
					(-1.096)	(-2.835)
FlowUniverse_ETF	0.000	0.000	**0.000*****	**0.000*****		
	(1.533)	(1.531)	(3.783)	(3.907)		
FlowUniverse_Cert					**0.341***	**0.343***
					(1.835)	(1.832)
R^2	0.19	0.14	0.43	0.42	0.11	0.10
Adj. R^2	0.16	0.13	0.41	0.41	0.08	0.08
Year dummies	Yes	yes	yes	yes	yes	yes
Index dummies	Yes	no	yes	no	yes	no

In Panel A, the table presents the results of the impact of flows on each other's flow estimated with one pooled ordinary least squares regression and two seemingly unrelated regressions. Panel B only presents the seemingly unrelated regressions. The sample includes 91 ETFs and 100 index certificates based on $\#_{Flow}$ that replicate 23 indices from 01/01/2007 to 12/31/2012. In Panel A, the dependent variables are aggregated monthly flows to ETFs (FlowETF) and index certificates (FlowCert) grouped by indices, whereas in Panel B the FlowETF is split into FlowETF_in and FlowETF_out. The independent variables are current and lagged flows to ETFs and index certificates, current and lagged returns, the logarithm of the market capitalization of all ETFs and index certificates, current total costs per share and the flow of the market volume of the universe of all ETFs and index certificates in Germany. Fixed effects are controlled with year and index dummies. The *t*-statistics are reported in parentheses. The symbols *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

4.2 Tracking Error and Tracking Difference

Total time period

Can the complementarity between ETFs and index certificates be explained by the tracking ability as well? Do both index products track their indices similarly, or do the index certificates, which compensate for the additional default risk, have higher tracking abilities than ETFs? Table 3, separated into two panels, presents the results of tracking errors and tracking differences. Panel A shows tracking abilities on a net of fees basis, whereas Panel B presents this information on a gross of fees basis. The difference in means between the two groups of index products is calculated for each index and tested for statistical significance.

The results show that both ETFs and index certificates suffer from high tracking errors. Moreover, ETFs and index certificates similarly replicate their underlying indices, thus explaining the complementarity between the two index products. The difference in the means of tracking errors between ETFs and index certificates is insignificant for all indices except for STOXX Europe 600 Industrial Goods & Services (both net and gross of fees) and STOXX Europe 600 (gross of fees).

Gross of fees, ETFs and index certificates underperform its indices on an average by 0.79% and 0.83%, respectively. Comparing the average return of ETFs and index certificates with their net total return indices, the underperformance is reduced to 0.12% in each case. Consequently, dividend taxation contributes 0.67% to the underperformance of ETFs and 0.71% to the underperformance of index certificates.

Net of fees, the difference in the means of tracking differences between ETFs and index certificates is statistically significant for five of the 23 indices. Index certificates mainly replicate these indices better than ETFs, except for the STOXX Europe 600 Automobiles & Parts Index. Gross of fees, the results are almost indistinguishable. The mean of tracking differences is significantly smaller for ETFs than for index certificates in three out of seven cases. The overall difference in the means of tracking differences deviates significantly from zero, preferring index certificates over ETFs.

Do high tracking errors imply high tracking differences? Figure A1 confirms the results of Table 3 indicating that tracking errors are much higher than tracking differences. Moreover, the tracking errors of index certificates differ from each other more strongly than the tracking errors of ETFs, whereas the tracking differences of index certificates do not vary as much as the tracking differences of ETFs. I also calculated the correlations between the tracking errors and tracking differences. They accordingly amounted to -0.05 for the full sample, -0.06 for ETFs and -0.05 for index certificates. All correlations are statistically insignificant.

Table 3: Tracking Error and Tracking Difference

	Tracking error						Tracking difference					
Panel A: Net total return indices	**ETFs**		**Certificates**		**Difference**		**ETFs**		**Certificates**		**Difference**	
EURO STOXX 50	0.0204	***	0.0229	***	-0.0025		0.0028	**	-0.0020	*	0.0048	***
EURO STOXX Select Dividend 30	0.0205	***	0.0271	**	-0.0066		0.0016		-0.0022		0.0038	
STOXX Europe 50	0.0216	***	0.0250	*	-0.0035		-0.0025		-0.0023		-0.0002	
STOXX Europe 600	0.0198	***	0.0171	**	0.0026		-0.0015		-0.0012		-0.0003	
STOXX Europe 600 Automobiles & Parts	0.0255	***	0.0284	***	-0.0029		0.0001		-0.0013		0.0015	*
STOXX Europe 600 Banks	0.0246	***	0.0280	***	-0.0033		-0.0032		-0.0019		-0.0014	
STOXX Europe 600 Basic Resources	0.0365	***	0.0325	***	0.0039		-0.0037	**	-0.0015	*	-0.0021	*
STOXX Europe 600 Chemicals	0.0199	**	0.0234	***	-0.0036		-0.0021		-0.0013	**	-0.0008	
STOXX Europe 600 Construction & Materials	0.0372		0.0241	**	0.0131		-0.0012		-0.0031	**	0.0019	
STOXX Europe 600 Financial Services	0.0220	*	0.0226	**	-0.0005		-0.0028		-0.0005		-0.0023	
STOXX Europe 600 Food & Beverage	0.0197	***	0.0234	**	-0.0036		-0.0046	*	-0.0009		-0.0037	**
STOXX Europe 600 Health Care	0.0199	***	0.0193	***	0.0006		-0.0033		-0.0013		-0.0020	
STOXX Europe 600 Industrial Goods & Services	0.0176	***	0.0161	***	0.0015	**	-0.0039		-0.0013		-0.0026	*
STOXX Europe 600 Insurance	0.0231	***	0.0249	***	-0.0018		-0.0025		-0.0009		-0.0016	
STOXX Europe 600 Media	0.0215	***	0.0210	***	0.0006		0.0009		-0.0004		0.0013	
STOXX Europe 600 Oil & Gas	0.0270	***	0.0238	***	0.0031		-0.0008		-0.0001		-0.0007	
STOXX Europe 600 Personal & Household Goods	0.0219	*	0.0204	**	0.0015		-0.0020		-0.0005		-0.0015	
STOXX Europe 600 Retail	0.0210	**	0.0190	**	0.0020		0.0001		-0.0001		0.0002	
STOXX Europe 600 Technology	0.0165	***	0.0163	**	0.0002		-0.0016		-0.0012	*	-0.0003	
STOXX Europe 600 Telecommunications	0.0234	***	0.0249	***	-0.0015		-0.0023		-0.0004		-0.0019	
STOXX Europe 600 Travel & Leisure	0.0254	**	0.0215	**	0.0039		-0.0027		-0.0003		-0.0024	
STOXX Europe 600 Utilities	0.0215	***	0.0233	**	-0.0019		0.0001		-0.0005		0.0006	
All net total return indices	0.0228	***	0.0234	***	-0.0006		-0.0012	***	-0.0012	***	0.0000	

Panel B: Gross total return indices	Tracking error						Tracking difference					
	ETFs		Certificates		Difference		ETFs		Certificates		Difference	
DAX 30	0.0202	***	0.0213	***	-0.0011		-0.0061	**	-0.0015	**	-0.0046	**
EURO STOXX 50	0.0211	***	0.0237	***	-0.0027		-0.0078	***	-0.0161	**	0.0083	*
EURO STOXX Select Dividend 30	0.0208	***	0.0281	**	-0.0072		-0.0118		-0.0191	**	0.0072	
STOXX Europe 50	0.0217	***	0.0252	*	-0.0035		-0.0087		-0.0089		0.0003	
STOXX Europe 600	0.0205	***	0.0172	**	0.0033	*	-0.0074	**	-0.0070	*	-0.0005	
STOXX Europe 600 Automobiles & Parts	0.0267	**	0.0290	***	-0.0024		-0.0047	**	-0.0065	***	0.0018	**
STOXX Europe 600 Banks	0.0249	***	0.0279	***	-0.0031		-0.0078	**	-0.0069	***	-0.0008	
STOXX Europe 600 Basic Resources	0.0328	***	0.0325	***	0.0003		-0.0053	**	-0.0034	**	-0.0019	*
STOXX Europe 600 Chemicals	0.0199	**	0.0250	***	-0.005		-0.0078	*	-0.0071	***	-0.0007	
STOXX Europe 600 Construction & Materials	0.0197	***	0.0248	**	-0.0051		-0.0106	**	-0.0115	***	0.0009	
STOXX Europe 600 Financial Services	0.0220	**	0.0231	**	-0.0011		-0.0083	*	-0.0062	***	-0.0022	
STOXX Europe 600 Food & Beverage	0.0222	**	0.0242	**	-0.0020		-0.0097	**	-0.0061	**	-0.0036	**
STOXX Europe 600 Health Care	0.0208	***	0.0209	***	-0.0001		-0.0098	**	-0.0084	***	-0.0014	
STOXX Europe 600 Industrial Goods & Services	0.0178	***	0.0164	***	0.0014	*	-0.0079	*	-0.0055	*	-0.0024	*
STOXX Europe 600 Insurance	0.0232	***	0.0261	***	-0.0029		-0.0083	*	-0.0070	**	-0.0013	
STOXX Europe 600 Media	0.0207	**	0.0223	***	-0.0016		-0.0051	*	-0.0075	**	0.0024	*
STOXX Europe 600 Oil & Gas	0.0258	**	0.0244	***	0.0014		-0.0070	*	-0.0086	*	0.0015	
STOXX Europe 600 Personal & Household Goods	0.0230	*	0.0205	**	0.0025		-0.0047		-0.0033	*	-0.0014	
STOXX Europe 600 Retail	0.0210	**	0.0194	**	0.0016		-0.0039		-0.0042	**	0.0002	
STOXX Europe 600 Technology	0.0176	**	0.0172	**	0.0004		-0.0066	*	-0.0066	**	0.0000	
STOXX Europe 600 Telecommunications	0.0259	**	0.0258	***	0.0001		-0.0119		-0.0130	**	0.0011	
STOXX Europe 600 Travel & Leisure	0.0243	**	0.0217	**	0.0025		-0.0047		-0.0025		-0.0023	
STOXX Europe 600 Utilities	0.0249	**	0.0240	***	0.0009		-0.0105	*	-0.0113	***	0.0009	
All gross total return indices	0.0224	***	0.0233	***	-0.0009		-0.0077	***	-0.0064	***	-0.0013	**
All gross total return indices excluding DAX 30	0.0226	***	0.0241	***	-0.0014		-0.0079	***	-0.0083	***	0.0005	

The table presents the means of annualized tracking errors and tracking differences on a net of fees (Panel A) and gross of fees basis (Panel B). Tracking error is calculated as the standard deviation of return differences between the ETF/index certificate and the corresponding index. Tracking difference measures the average difference of returns between the ETF/index certificate and their index. The difference in means between both ETFs and index certificates is calculated for each index and tested for statistical significance. The sample is based on $\#_{TE}$ and consists of 72 ETFs and 113 index certificates replicating 23 indices from 06/01/2009 to 12/31/2012. The symbols *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Rolling tracking error and tracking difference

Do tracking errors and tracking differences change over time? If this is true, then can we identify time periods in which ETFs are superior to index certificates or the other way around? To answer these questions, the whole examination period is divided into 660 one-year periods from 06/01/2009 to 12/31/2011. The tracking errors and the tracking differences of ETFs and index certificates are calculated for each sub-period and presented in two ways. Figure A2 shows the means of rolling tracking errors of ETFs and index certificates, whereas Figure A3 illustrates the means of rolling tracking differences. Moreover, Table 4 presents the percentages of significant statistical differences in tracking errors and tracking differences between ETFs and index certificates for each index over time.

The results in Figure A2 suggest that the means of tracking errors change over time. However, the differences in means of tracking errors between the two groups scarcely differ from each other. Table 4 shows that the overall difference in tracking errors is only statistically significant in 13% of all time periods. Moreover, only ETFs exhibit significantly smaller tracking errors which are limited to the period from June 2011 to September 2011. Nevertheless, tracking errors of ETFs and index certificates are always significantly different from zero.

Based on the indices, either tracking errors of ETFs are only significantly smaller than tracking errors of index certificates over time or it is the other way round. ETFs track two of the 23 indices better than index certificates, namely DAX 30 and EURO STOXX 50. Table 4 shows that this applies to 42% and 34% of all the 660 sub-periods, respectively. Index certificates exhibit significantly smaller tracking errors for five indices (e.g. STOXX Europe 600 Media and STOXX Europe 600 Oil & Gas).

Compared to tracking errors, tracking differences indicate different results. Figure A3 shows that the average tracking differences of ETFs and index certificates similarly fluctuate for the rolling time periods. Nonetheless, the differences in the tracking differences between ETFs and index certificates are statistically significant in 69% of all time periods as shown in Table 4. Based on the indices, Table 4 also shows that ETFs replicating EURO STOXX 50 exhibit higher tracking differences than index certificates. In comparison, index certificates replicating DAX 30 and STOXX Europe 600 Basic Resources have higher tracking differences than ETFs. Interestingly, neither ETFs nor index certificates succeed in producing higher returns than the underlying index except for ETFs replicating EURO STOXX 50. ETFs and index certificates underperform its indices in 65% and 69% of all the 660 time periods, respectively.

Table 4: Rolling Tracking Error and Tracing Difference: Percentage of Significant Group Differences

	Difference in tracking errors			Difference in tracking differences			Tracking difference			
	Total	ETF < Cert	ETF > Cert	Total	ETF < Cert	ETF > Cert	ETF < Index	ETF > Index	Cert < Index	Cert > Index
DAX 30	0.42	0.42	0	0.93	0.93	0	0.93	0	0.64	0.16
EURO STOXX 50	0.34	0.34	0	0.63	0	0.63	0.01	0.55	0.25	0.02
EURO STOXX Select Dividend 30	0	0	0	0.02	0	0.02	0	0	0.11	0
STOXX Europe 50	0.09	0.09	0	0	0	0	-	-	-	-
STOXX Europe 600	0	0	0	0.12	0.03	0.10	0.15	0.07	-	-
STOXX Europe 600 Automobiles & Parts	0	0	0	0.43	0.18	0.25	-	-	0.33	0.07
STOXX Europe 600 Banks	0	0	0	0.11	0.07	0.04	0.13	0	0.39	0.05
STOXX Europe 600 Basic Resources	0.40	0	0.40	0.60	0.60	0.01	0.52	0	0.47	0.09
STOXX Europe 600 Chemicals	0	0	0	0.05	0.01	0.04	-	-	0.31	0.13
STOXX Europe 600 Construction & Materials	0	0	0	0.23	0.01	0.22	-	-	0.51	0.04
STOXX Europe 600 Financial Services	0	0	0	0.14	0.03	0.11	-	-	0.27	0.10
STOXX Europe 600 Food & Beverage	0	0	0	0.43	0.43	0	0.21	0	0.24	0.16
STOXX Europe 600 Health Care	0.24	0	0.24	0.16	0.16	0	0.10	0	0.23	0.11
STOXX Europe 600 Industrial Goods & Services	0.04	0	0.04	0.08	0.08	0	0.08	0	0.08	0.02
STOXX Europe 600 Insurance	0.02	0.02	0	0.04	0.01	0.03	0.01	0	0.18	0.09
STOXX Europe 600 Media	0.60	0	0.60	0.19	0.02	0.17	-	-	0.22	0.08
STOXX Europe 600 Oil & Gas	0.65	0	0.65	0.09	0.08	0.01	0.09	0.01	0.26	0.08
STOXX Europe 600 Personal & Household Goods	0	0	0	0.17	0.17	0	-	-	0.17	0.09
STOXX Europe 600 Retail	0	0	0	0.09	0.06	0.03	-	-	0.23	0.09
STOXX Europe 600 Technology	0.01	0	0.01	0.06	0.05	0.01	0.05	0	0.20	0.02
STOXX Europe 600 Telecommunications	0.06	0	0.06	0.01	0	0.01	0	0	0.25	0.09
STOXX Europe 600 Travel & Leisure	0.22	0	0.22	0.11	0.11	0	-	-	0.14	0.05
STOXX Europe 600 Utilities	0	0	0	0.01	0	0.01	0	0	0.31	0.13
All indices	0.13	0.13	0	0.69	0.44	0.26	0.65	0.10	0.69	

The table presents the percentages of significant statistical difference in tracking errors and tracking differences between ETFs and index certificates for each index. The table also shows the percentage of significant negative and positive tracking differences against the underlying index. Analyses are based on $\#_{TE}$ and 660 periods of time. Differences are counted as significant, if t-statistics indicate statistical significance at less than the 10% level.

Summing up, the results have slightly different implications depending on tracking error or tracking difference. For this reason, we can ask for the best suited measurement to evaluate the replication ability of ETFs and index certificates. The tracking error measures both the positive and negative return differences between the index product and the underlying index, whereas negative return differences are offset by positive return differences using the tracking difference. Because index investors expect the return of the underlying index at any time, the tracking error is a more important factor in this case. Based on the analyses, ETFs and index certificates, on an average, replicate their indices in the same way for the total time period and rolling time periods, indicating the complementarity between the two index products. For example, only individual indices like DAX 30 and EURO STOXX 50 are better tracked by ETFs, whereas STOXX Europe 600 Media and STOXX Europe 600 Oil & Gas are better tracked by index certificates over time.

4.3 Factors Influencing Tracking Errors

Which factors affect tracking errors? Can we identify differences in factors between ETFs and index certificates as well as differences in influence over time? Table 5 illustrates the results of the influencing factors on the tracking error of ETFs and index certificates using a cross-sectional regression, a pooled-cross-sectional regression as well as a panel-regression with fixed-effects. In Panel A, the tracking errors as the dependent variable are calculated individually for each ETF and index certificate, whereas Panel B shows the results depending on the means of tracking errors of each index.

The results indicate that the influence of the analyzed factors differ between regressions. But which regression is best suited to analyze the influencing factors? The panel-regression with fixed-effects is the most appropriate one because this regression takes into consideration the differences between ETFs and index certificates as well as changes over time. However, all three regressions can explain between 50% and 94% of the variation in tracking errors.

The analyzed factors influence the tracking errors of ETFs and index certificates differently. Based on the fixed-effects regression, tracking errors of ETFs are only positively related to dividends in Panel A and B. Spread and volume affect tracking errors positively in Panel A and negatively in Panel B, whereas risk and securities lending influence tracking errors negatively in Panel A and positively in Panel B. Interestingly, costs only influence the tracking errors of individual ETFs positively. A possible reason for this is that the impact of costs is either a less important factor based on indices or that it is compensated more strongly by other factors. For index certificates, the impact of spread (Panel A) and risk (Panel B) on tracking errors is positive. Volume yields dif-

ferent results, affecting the tracking errors of index certificates negatively in Panel A and positively in Panel B.

The risk factor influenced the tracking ability of ETFs and index certificates more strongly after July 2011 because investors' aversion to risks increased dramatically when the financial crisis started to spread to Italy and Spain after entering Greece. As a result, the investment activity in both ETFs and index certificates decreased after July 2011. Decreasing liquidity increases the spread of index products and therefore increases tracking errors. Furthermore, the influence of securities lending on tracking errors turns positive for individual ETFs. This is not surprising because physical ETFs enhance their engagement in securities lending to improve their returns and to compensate for outflows due to transaction costs.

Table 5: Influencing Factors: Product-, Index- and Time-Dependent Analyses

	Cross-sectional regression		**Pooled-cross-sectional regression**		**Fixed-effects regression**	
Panel A	**ETF**	**Cert**	**ETF**	**Cert**	**ETF**	**Cert**
Intercept	**0.0003*** **	**0.0005*** **	0.0003	**0.0006*** **	**0.0011*** **	**0.0009*** **
	(18.845)	(16.587)	(1.198)	(4.128)	(38.196)	(9.389)
Risk	**0.0453*** **	**0.0702*** **	**0.0433*** **	**0.0675*** **	**-0.0235*** **	0.0081
	(56.579)	(26.809)	(4.827)	(13.688)	(-14.819)	(1.332)
Spread	**0.0079*** **	**0.0097*** **	0.0136	**0.0103** **	**0.0178*** **	**0.0095* **
	(6.662)	(4.853)	(0.749)	(2.384)	(17.554)	(1.868)
Volume	0.0000	**0.0000*** **	0.0000	0.0000	**0.0000*** **	**0.0000*** **
	(-0.005)	(-3.488)	(0.397)	(-0.829)	(10.296)	(-2.590)
Dividends	**0.1522*** **		**0.1572*** **		**0.0937*** **	
	(38.645)		(2.945)		(23.170)	
Costs	**1.0306*** **		**1.0597*** **		**0.6244*** **	
	(68.063)		(6.446)		(13.261)	
SecLending	**-0.0711*** **		-0.0706		**-0.0792*** **	
	(-19.505)		(-1.148)		(-5.914)	
Risk*Dummy	**-0.0046*** **	**-0.0205*** **	-0.0059	**-0.0305** **	**0.0191*** **	**0.0335*** **
	(-5.083)	(-8.425)	(-0.633)	(-2.397)	(26.446)	(5.409)
Spread*Dummy	**0.0364*** **	**0.0611*** **	0.0278	**0.0574*** **	**0.0039*** **	**0.0430*** **
	(23.795)	(15.765)	(1.534)	(8.860)	(4.359)	(22.824)
Volume*Dummy	**0.0000*** **	**0.0000*** **	0.0000	0.0000	**0.0000** **	**0.0000*** **
	(3.114)	(3.445)	(-0.314)	(0.543)	(-2.157)	(7.667)
Dividends*Dummy	**-0.0729*** **		-0.0622		**-0.0654*** **	
	(-14.413)		(-1.082)		(-24.263)	
Costs*Dummy	**-0.1503*** **		-0.1849		**-0.2266*** **	
	(-7.874)		(-1.076)		(-15.346)	
SecLending*Dummy	**0.0855*** **		0.0828		**0.0861*** **	
	(21.817)		(1.402)		(37.845)	
Adj. R^2	0.50	0.57	0.50	0.57	0.84	0.82
Product dummies	no	no	no	no	Yes	yes
Time dummies	no	no	yes	yes	Yes	yes

Panel B						
Intercept	0.0001	**0.0010*** **	0.0000	**0.0012*** **	0.0003	**0.0012*** **
	(1.320)	(16.010)	(0.000)	(4.519)	(1.050)	(12.629)
Risk	**0.0400*** **	**0.0472*** **	**-0.0008** **	**0.0487*** **	**0.0739*** **	**0.0111** **
	(18.275)	(20.029)	(-2.223)	(5.112)	(7.024)	(2.165)
Spread	**-0.0127*** **	**0.0318*** **	**0.0594** **	**0.0398*** **	**-0.0225*** **	0.0054
	(-4.927)	(11.075)	(2.228)	(4.507)	(-5.109)	(1.114)
Volume	**0.0000*** **	**0.0000*** **	-0.0248	0.0000	**0.0000*** **	**0.0000*** **
	(-7.800)	(-6.376)	(-1.268)	(-0.627)	(-5.116)	(3.381)
Dividends	**0.5214*** **		**0.0000* **		**0.8727*** **	
	(12.511)		(-1.890)		(12.871)	
Costs	**2.5253*** **		**0.7404** **		-0.7683	
	(13.196)		(2.445)		(-1.027)	
SecLending	0.0209		**1.8935*** **		**0.3079*** **	
	(0.673)		(5.858)		(3.490)	
Risk*Dummy	**0.0224*** **	**-0.0064** **	**-0.1942* **	**-0.0339*** **	**0.0390*** **	-0.0008
	(9.193)	(-2.349)	(-1.754)	(-2.596)	(2.934)	(-0.217)
Spread*Dummy	-0.0004	**-0.0082** **	0.0502	**-0.0251** **	**0.0269*** **	**-0.0376*** **
	(-0.228)	(-2.506)	(0.972)	(-2.176)	(4.245)	(-11.182)
Volume*Dummy	**0.0000* **	**0.0000*** **	**0.0435** **	0.0000	**0.0001*** **	**0.0000*** **
	(1.783)	(2.897)	(2.043)	(-1.007)	(6.384)	(-2.811)
Dividends*Dummy	0.0251		0.0000		**-0.6074*** **	
	(0.526)		(0.193)		(-7.401)	
Costs*Dummy	-0.1853		-0.1440		**-1.8600*** **	
	(-1.300)		(-1.000)		(-3.295)	
SecLending*Dummy	**-0.1978*** **		**2.4281** **		**-0.5125*** **	
	(-7.937)		(2.134)		(-5.017)	
Adj. R^2	0.92	0.59	0.93	0.55	0.94	0.88
Product dummies	no	no	no	no	Yes	yes
Time dummies	no	no	yes	yes	Yes	yes

The table presents the results of factors influencing tracking errors based on one cross-sectional regression and two panel-regressions (pooled-cross-sectional and fixed-effects). The number of ETFs and index certificates is based on $\#_{TE}$. In Panel A, tracking errors are calculated as the one-year standard deviation of return differences for each ETF and certificate *i* at time *t*. In Panel B, the analysis is repeated with the mean of tracking errors based on index *i* at time *t*. Fixed effects are controlled by including dummies based on time and the index products (ETFs and index certificates) as well as the indices. Explanatory variables are: risk, which is the standard deviation of daily returns; spread, which is the average of daily relative spreads; volume, which is the natural logarithm of the average daily trading volume; dividends, as the ratio of average dividends and average trading price; securities lending, as the ratio of net 12 month securities lending revenue and average assets under management over the same period; total costs, as the average daily total costs per share. Changes of influence in explanatory variables after July 2011 are controlled by including one dummy variable. The *t*-statistics are reported in parentheses. The symbols *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Another explanation for the increase of tracking errors could be that the absolute value of factors influencing tracking errors changed after July 2011. Factors positively influencing tracking errors should increase in absolute terms, whereas factors with negative impact on tracking errors should decrease. Table 6 illustrates the results of panel-regressions controlled by fixed effects for individual index products (Panel A) and indices (Panel B), respectively.

Before July 2011, the level of risk of ETFs was comparable to the risk level of index certificates. The spread of index certificates is nearly twice as high as the spread of ETFs. In comparison, the volume of index certificates is lower than the volume of ETFs. After July 2011, all analyzed influ-

encing factors increased significantly except for dividends and securities lending of ETFs in Panel A as well as the spread of index certificates in both Panel A and B.

Table 6 supports the fact that risk and spread are the main drivers that increased the tracking errors of ETFs after July 2011. According to Table 5, the impacts of risk and spread on tracking errors increase or become positive. The increase in the tracking errors of index certificates can mainly be explained by an increase in risk along with an increase in the impact of risk (Table 5, Panel A). All other analyzed factors yield mixed results since either the impact increases and the absolute value decreases or vice versa.

Table 6: Influencing Factors before and after July 2011

	ETFs			Certificates		
Panel A: Product based	**Intercept**	**Dummy**	**Adj. R^2**	**Intercept**	**Dummy**	**Adj. R^2**
Risk	**0.0128*** ** (150.228)	**0.0021*** ** (18.524)	0.76	**0.0126*** ** (150.708)	**0.0022*** ** (19.717)	0.72
Spread	**0.0026*** ** (421.775)	**0.0003*** ** (31.863)	0.94	**0.0051*** ** (3482.906)	**0.0000*** ** (-7.623)	1.00
Volume	**13.4375*** ** (3479.254)	**0.0349*** ** (6.755)	0.98	**10.7320*** ** (1244.335)	**0.1545*** ** (13.143)	0.89
Dividends	**0.0003*** ** (4404.795)	**0.0000*** ** (-5.185)	0.88			
Costs	**0.0004*** ** (851.427)	**0.0000*** ** (43.027)	0.98			
Securities lending	**0.0002*** ** (2244709)	**0.0000*** ** (-16.777)	1.00			
Panel B: Index based						
Risk	**0.0127*** ** (152.935)	**0.0018*** ** (16.714)	0.76	**0.0128*** ** (161.423)	**0.0014*** ** (12.880)	0.77
Spread	**0.0027*** ** (370.538)	**0.0004*** ** (37.541)	0.90	**0.0062*** ** (283.601)	**-0.0004*** ** (-14.363)	0.93
Volume	**12.9633*** ** (2194.216)	**0.1167*** ** (14.727)	0.97	**9.9032*** ** (693.363)	**0.1731*** ** (8.562)	0.89
Dividends	**0.0009*** ** (447.807)	**0.0001*** ** (21.557)	0.98			
Costs	**0.0004*** ** (805.369)	**0.0000*** ** (43.827)	0.98			
Securities lending	**0.0015*** ** (2036.497)	**0.0001*** ** (51.231)	1.00			

The table presents the results of changes in the influencing factors of both ETFs and index certificates based on $\#_{TE}$ using panel-regressions. Fixed effects are controlled by dummies for index products in Panel A and dummies for indices in Panel B. Dependent variables are the respective means of influencing factors over time including risk, spread, volume, dividends, securities lending and total costs. The independent variable consists of one dummy variable, separating the examination period into periods before and after July 2011. *t*-statistics are reported in parentheses. The symbols *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

5 Conclusion

This study contributes to the literature by examining the competition between ETFs and index certificates within one market. One might assume that ETFs have started to replace index certifi-

cates, especially after their loss of reputation during the collapse of Lehman Brothers. On the contrary, this study finds that ETFs and index certificates are complements, but not perfect complements. An increase in flows to ETFs results in an increase in flows to index certificates and vice versa. Based on this result, this study finds similar tracking abilities and a clientele effect, which separates investors into different market niches, to vouch for the coexistence of both index products in one market.

Moreover, this study analyzes the tracking abilities of ETFs and index certificates over time and discovers explanations for possible changes in replication quality. This study finds that both tracking errors and tracking differences change significantly over time. However, the differences in the means of tracking errors between ETFs and index certificates scarcely differ from each other. Only individual indices such as DAX 30 and EURO STOXX 50 are better tracked by ETFs, whereas STOXX Europe 600 Media and the STOXX Europe 600 Oil & Gas are better tracked by index certificates over time, for example. Indeed, the differences in the tracking differences between ETFs and index certificates are statistically significant in 69% of all time periods. Both index products mainly underperform its indices.

Changes in the tracking ability of ETFs over time can be explained by an increase in the impact and the absolute value of both risk and spread. The tracking ability of individual index certificates is primarily induced by risk. Therefore, investment decisions should depend on the current level of tracking ability as well as the current level of influencing factors as compared to the past.

References

Agapova, A. (2011). Conventional mutual index funds versus exchange-traded funds. Journal of Financial Markets 14 (2), pp. 323-343.

BlackRock (2012). ETP landscape. Industry highlights. BlackRock Inc., December 2012.

Blitz, D., Huij, J., Swinkels, L. (2012). The performance of European index funds and exchange-traded funds. European Financial Management 18 (4), pp. 649-662.

Elton, E. J., Gruber, M. J., Comer, G., Li, K. (2002). Spiders: where are the bugs? The Journal of Business 75 (3), pp. 453-472.

Frino, A., Gallagher, D. R. (2001). Tracking S&P 500 index funds. The Journal of Portfolio Management 28 (1), pp. 44-55.

Gastineau, G. L. (2004). The benchmark index ETF performance problem. A simple solution. The Journal of Portfolio Management 30 (2), pp. 96-103.

Graf, S. (2001). Eins zu Eins mit dem Index? Working paper, Schweizerisches Institut für Banken und Finanzen, November.

Johanning, L., Becker, M., Seeber, M. (2011). Unterschiede und Gemeinsamkeiten börsennotierter, passiver Investmentprodukte. Discussion paper, EDA/vwd academy AG, October.

Johnson, B., Bioy, H., Kellett, A., Davidson, L. (2013). On the right track: measuring tracking efficiency in ETFs. Morningstar ETF Research, February.

Klein, C., Kundisch, D. (2008). Indexzertifikat oder ETF? Eine entscheidungstheoretische Analyse. Zeitschrift für Planung & Unternehmenssteuerung 19 (3), pp. 353-370.

Klein, C., Kundisch, D. (2009a). Der Tracking Error von indexnachbildenden Instrumenten auf den DAX – eine empirische Analyse des ETFs DAX EX sowie zehn Indexzertifikaten. Der Betrieb 62 (22), pp. 1141-1145.

Klein, C., Kundisch, D. (2009b). Zur Preissetzung verschiedener Emittenten bei Indexzertifikaten auf den DAX. Zeitschrift für Bankrecht und Bankwirtschaft 21 (3), pp. 212-224.

Kostovetsky, L. (2003). Index mutual funds and exchange-traded funds. The Journal of Portfolio Management 29 (4), pp. 80-92.

Meinhardt, C., Mueller, S., Schoene, S. (2015). Physical and synthetic exchange-traded funds: the good, the bad, or the ugly? Journal of Investing 24 (2), pp. 35-44.

Milonas, N. T., Rompotis, G. G. (2006). Investigating European ETFs: The case of the Swiss exchange traded funds. Working paper, presented at the annual conference of the HFAA in Thessaloniki.

Pope, P. F., Yadav, P. K. (1994). Discovering errors in tracking error. The Journal of Portfolio Management 20 (2), pp. 27-32.

Poterba, J. M., Shoven, J. B. (2002). Exchange-traded funds: a new investment option for taxable investors. The American Economic Review 92 (2), pp. 422-427.

Rompotis, G. G. (2005). An empirical comparing investigation on exchange traded funds and index funds performance. Working paper, December.

Schmidhammer, C., Lobe, S., Röder, K., (2011). Intraday pricing of ETFs and certificates replicating the German DAX index. Review of Managerial Science 5 (4), pp. 337-351.

Shin, S., Soydemir, G. (2010). Exchange-traded funds, persistence in tracking errors and information dissemination. Journal of Multinational Financial Management 20 (4-5), pp. 214-234.

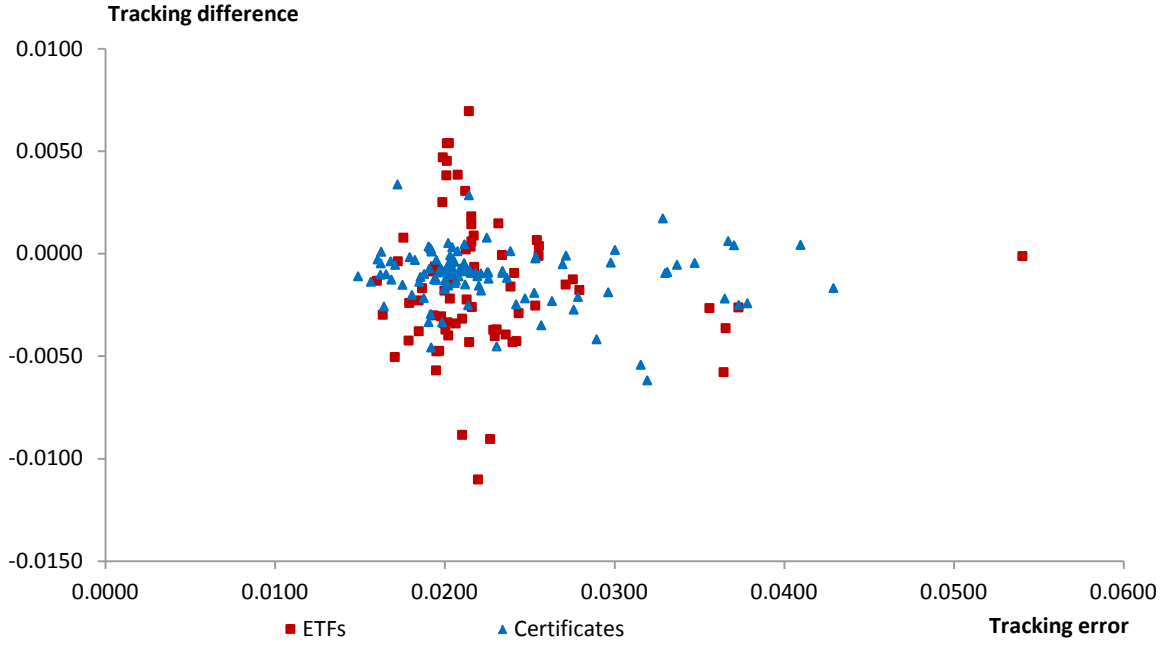
Sirri, E. R., Tufano, P. (1998). Costly search and mutual fund flows. The Journal of Finance 53 (5), pp. 1589-1622.

Svetina, M., Wahal, S. (2008). Exchange traded funds: Performance and competition. Working paper, November.

Yan, A. (2006). Leasing and debt financing: substitutes or complements? Journal of Financial and Quantitative Analysis 41 (3), pp. 709-731.

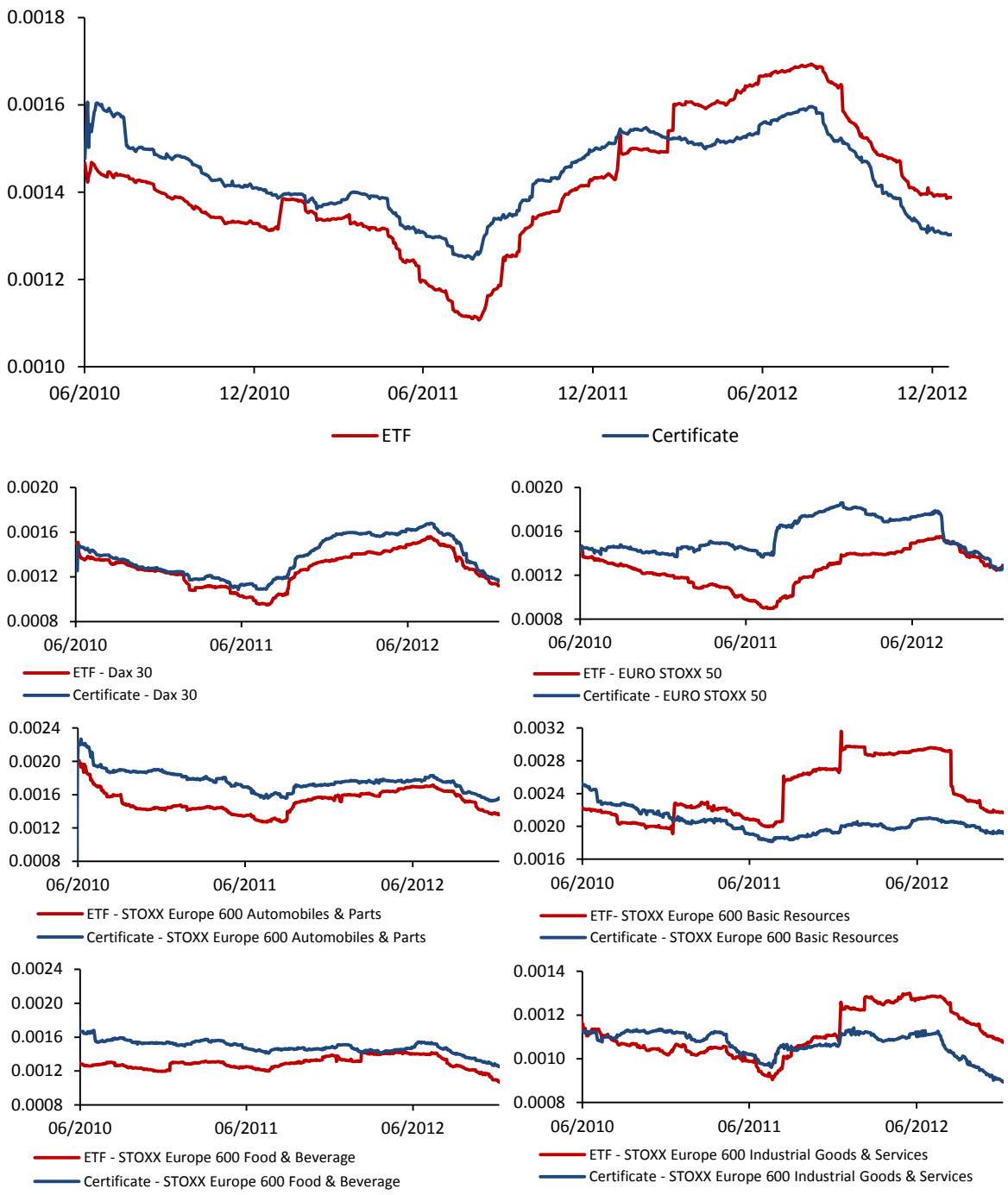
Appendix

Figure A1: Tracking Error vs. Tracking Difference



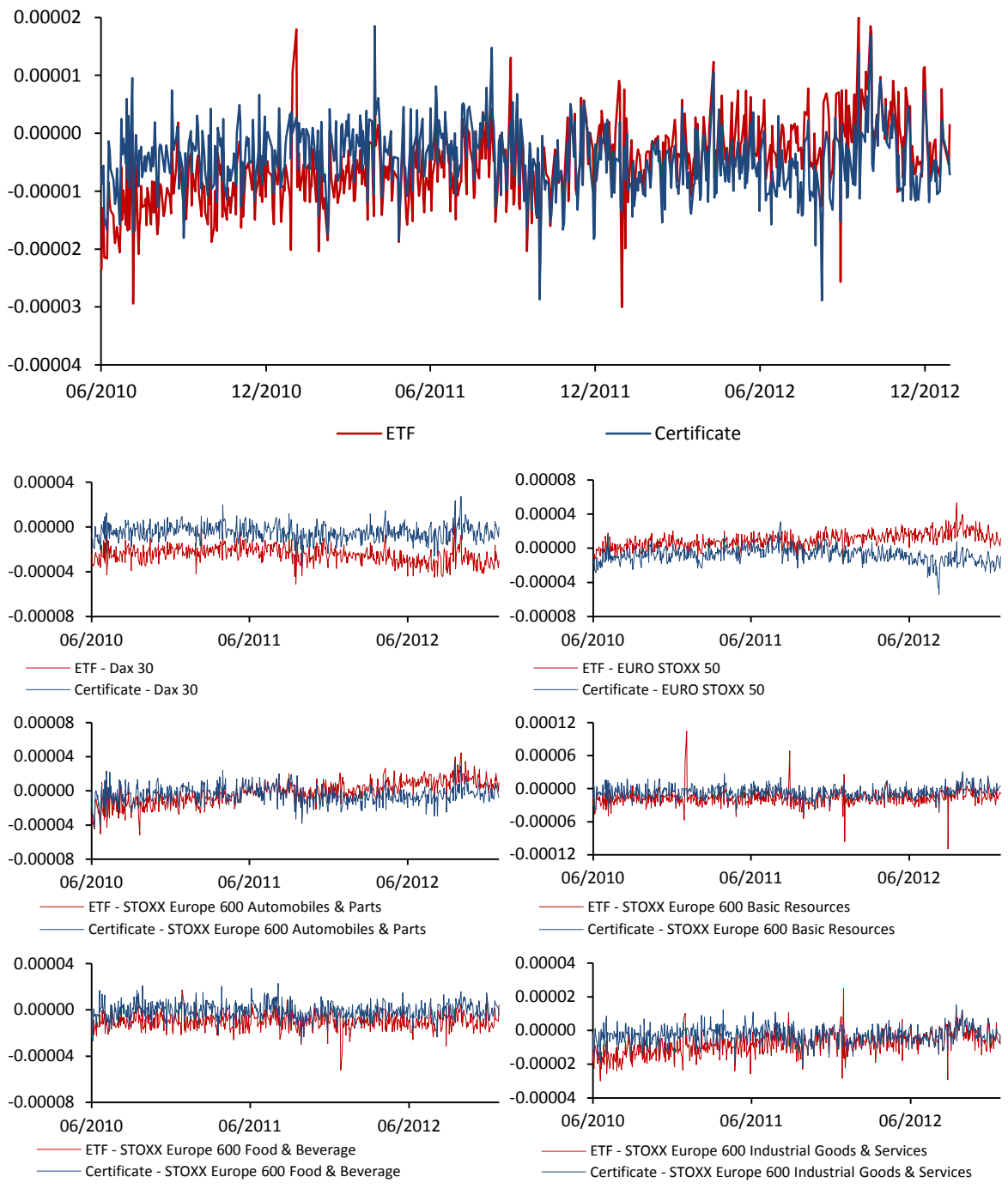
The figure compares the annualized tracking error of each ETF and index certificate with the annualized tracking difference for the total time period from 06/01/2009 to 12/31/2012.

Figure A2: Rolling Tracking Error: ETFs vs. Index Certificates



The figure presents the means of daily rolling one-year tracking errors of ETFs and index certificates for the overall sample and six selected indices. The overall sample includes 72 ETFs and 113 certificates from 06/01/2009 to 12/31/2012.

Figure A3: Rolling Tracking Difference: ETFs vs. Index Certificates



The figure presents the means of daily rolling one-year tracking differences of ETFs and index certificates for the overall sample and six selected indices. The overall sample includes 72 ETFs and 113 certificates from 06/01/2009 to 12/31/2012.

Part III:
Fondsbewertungen in Deutschland – Ein Vergleich der Vorhersagekraft

Abstract

Diese Arbeit geht der Frage nach, ob Fondsbewertungen wie das Feri Trust Rating, die Finanztest-Bewertung und die FondsNote in der Lage sind, die zukünftige Performance von Fonds vorherzusagen. Zudem wird untersucht, welche Faktoren die Vorhersagekraft beeinflussen. Hierzu zählen die Bewertungsmethodik, das Investitionsverhalten und die Fondskosten.

Der Vergleich der drei Fondsbewertungen zeigt, dass sich die Vorhersagekraft unterscheidet. Die FondsNote, die neben quantitativen auch qualitative Faktoren berücksichtigt, besitzt die höchste Vorhersagekraft. Den anderen beiden rein quantitativen Fondsbewertungen gelingt die Prognose der zukünftigen Fondsperformance deutlich schlechter. Zudem wird gezeigt, dass sich aufgrund einer Fondsbewertung das Verhalten der Investoren ändert. Dieses hat wiederum einen Einfluss auf die Vorhersagekraft der Fondsbewertung. Daneben stellen die Fondskosten einen weiteren wichtigen Einflussfaktor dar.

JEL-Klassifizierung: G1, G11, G12, G14

Schlüsselwörter: Fondsbewertung, Rating, Ranking, Prognosefähigkeit, Investitionsverhalten

Für wertvolle Hinweise und Anregungen danke ich Frau Prof. Dr. Sigrid Müller und Herrn Dr. Stefan Schöne vom Institut für Finanzierung. Feri Trust, Finanztest und Axel Springer/FondsConsult stellten mir dankenswerterweise Ihre monatlichen Fondsbewertungen zur Verfügung.

1 Einleitung

Der Markt für Fonds hat sich in den letzten Jahren stürmisch entwickelt. Ende 1991 waren laut dem Bundesverband Investment und Asset Management (BVI) insgesamt 381 Publikumsfonds in Deutschland zum Vertrieb zugelassen. Bis Ende 2013 stieg ihre Zahl auf 7.611. Ein Investor steht somit vor einer schwierigen Auswahlentscheidung. Anbieter von Fondsbewertungen wie Morningstar, Standard & Poor's, Feri Trust, Axel Springer/FondsConsult und auch die Zeitschrift Finanztest wecken mit ihren Einstufungen die Erwartung, bei der Fondsauswahl hilfreich zu sein. Hieraus lassen sich zwei Fragestellungen ableiten, die in wissenschaftlichen Arbeiten häufig thematisiert wurden.

1. Sind Fondsbewertungen in der Lage, die zukünftige Wertentwicklung eines Fonds vorherzusagen? (siehe z. B.: Blake und Morey (2000), Morey und Gottesman (2006), Duret et al. (2008) und Antypas et al. (2009))
2. Hat die Einstufung der Fonds durch Fondsbewertungen einen Einfluss auf das Investitionsverhalten der Investoren? (siehe z. B.: Sirri und Tufano (1998), Del Guercio und Tkac (2008) und Füss et al. (2010))

Im Mittelpunkt der beiden Fragestellungen stand bis dato vor allem das Morningstar Rating. Investoren fokussieren sich allerdings bei ihrer Fondsauswahl nicht ausschließlich auf diese eine Fondsbewertung. Sie können zwischen verschiedenen Fondsbewertungen wählen. Es stellt sich daher die Frage, ob die Ergebnisse zum Morningstar Rating auch für andere Fondsbewertungen gelten. Aus diesem Grund wird in dieser Arbeit neben dem Feri Trust Rating auch die Finanztest-Bewertung und die FondsNote in Bezug auf beide Fragestellungen untersucht. Diese Arbeit ist somit die Erste, die gleichzeitig drei für Investoren alternative Fondsbewertung in den Fokus der Untersuchung stellt.

Analysiert wird das Feri Trust Rating, da es sich nach dem Morningstar Rating um die zweitgrößte Fondsbewertung in Europa handelt. Die Finanztest-Bewertung besitzt eine hohe Reputation bei Investoren. Sie wird von Stiftung Warentest, der bekanntesten und größten Verbraucherschutzorganisation in Deutschland, herausgegeben. Beide Fondsbewertungen beruhen, ebenso wie das Morningstar Rating, ausschließlich auf quantitativen Faktoren. Die FondsNote hingegen berücksichtigt in ihrer Methodik neben quantitativen auch qualitative Faktoren wie die Managerkontinuität und den Investmentansatz. Dies ermöglicht die Untersuchung, ob qualitative Faktoren die Vorhersagekraft der Wertentwicklung eines Fonds beeinflussen, vielleicht sogar verbessern.

Der Vergleich des Feri Trust Ratings mit der Finanztest-Bewertung und der FondsNote zeigt, dass sich die Vorhersagekraft der drei Fondsbewertungen unterscheidet. Die Vorhersagekraft der FondsNote ist tatsächlich am höchsten. Den beiden rein quantitativen Fondsbewertungen gelingt die Prognose der zukünftigen Fondsperformance deutlich schlechter. Zudem zeigt auch der Vergleich der FondsNote mit Fondsbewertungen, die auf dem 4-Faktor-Alpha und dem geometrischen Mittel basieren, dass qualitative Faktoren zu einer Verbesserung der Vorhersagekraft führen. Im Gegensatz hierzu ergeben diese alternativen Fondsbewertungen eine bessere Differenzierung von sich zukünftig besser und schlechter entwickelnden Fonds als mit der Finanztest-Bewertung. Zudem wird in dieser Arbeit gezeigt, dass sich aufgrund einer Fondsbewertung das Investitionsverhalten ändert. Die hiermit verbundene Veränderung des Fondsvermögens übt wiederum einen Einfluss auf die Vorhersagekraft einer Fondsbewertung aus. Letztlich stellen die Fondskosten einen weiteren untersuchten Einflussfaktor dar.

Die Arbeit ist in sechs Kapitel unterteilt. Kapitel 2 gibt einen Überblick wissenschaftlicher Arbeiten, die sich mit Fondsbewertungen befassen. In Kapitel 3 werden die drei zu untersuchenden Fondsbewertungen anhand ihrer Methodik gegenübergestellt und Gemeinsamkeiten beziehungsweise Unterschiede herausgearbeitet. Kapitel 4 befasst sich mit dem Datensatz und der verwendeten Methodik. Kapitel 5 stellt die Ergebnisse vor und Kapitel 6 fasst diese kurz zusammen.

2 Literaturüberblick

Die Literatur beschränkt sich fast ausschließlich auf Arbeiten zum Morningstar Rating. Blake und Morey (2000) sind die Ersten, die die Vorhersagekraft des Morningstar Ratings in Bezug auf die künftige Wertentwicklung eines Fonds untersuchen. Sie zeigen für amerikanische Aktienfonds, dass die am zweithöchsten Bewerteten sich nicht in ihrer Performance von jenen mit Top-Bewertung unterscheiden. Sehr niedrig bewertete Aktienfonds weisen hingegen häufiger eine geringere zukünftige Performance auf. Zu sehr ähnlichen Ergebnissen gelangen Kräussl und Sandelowsky (2007), Antypas et al. (2009) und Füss et al. (2010). Kräussl und Sandelowsky (2007) verwenden einen Datensatz von Fonds, der das gesamte amerikanische Anlageuniversum abbildet. Antypas et al. (2009) untersuchen sämtliche USD Aktienfonds mit Morningstar Rating. Füss et al. (2010) analysieren die Vorhersagekraft des Morningstar Ratings aller in Deutschland zum Vertrieb zugelassenen Fonds.

Gerrans (2006) ist der Erste, der den Einfluss quantitativer und qualitativer Faktoren auf die Vorhersagekraft des Morningstar Ratings untersucht. Dies ist möglich, weil das Rating von Mor-

ningstar in Australien bereits von 1999 bis 2005 qualitative Faktoren berücksichtigt. Gerrans (2006) zeigt allerdings, dass sich die geringe Vorhersagekraft nicht verbessert.

Morey und Gottesman (2006) kommen in Bezug auf das Morningstar Rating zu anderen Ergebnissen. Sie schlussfolgern, dass dem in 2002 angepassten Morningstar Rating die Performanceprognose amerikanischer Aktienfonds sehr viel besser gelingt. Aktienfonds mit höheren Ratings weisen demnach gegenüber niedriger Bewerteten eine höhere zukünftige Performance auf. Müller und Weber (2014) untersuchen sämtliche in Deutschland zugelassene Fonds mit Finanztest-Bewertung. Die Ergebnisse zeigen, dass diese Bewertung eine Vorhersagekraft besitzt. Die Vorhersagekraft bezüglich deutscher Aktienfonds ist jedoch gering.

Garnier und Pujol (2007), Duret et al. (2008) und Hereil et al. (2010) untersuchen im Vergleich zu den vorherigen Arbeiten nicht die zukünftige Performance nach einer Fondsbewertung, sondern die Stabilität einer einmal vergebenen Bewertung im Zeitablauf. Hierbei wird gezeigt, dass Fonds ihr Morningstar Rating nur über sehr kurzfristige Zeiträume beibehalten. Beispielsweise erhalten jene Fonds Top-Bewertungen, die in der Vergangenheit im Vergleich zur Peer Group eine überproportionale Wertentwicklung verzeichneten. Allerdings behalten diese Fonds ihre hohe Performance nur selten bei, was zu einer Morningstar-Herabstufung innerhalb kürzester Zeit führt.

Ob Fondsbewertungen einen Einfluss auf das Investitionsverhalten haben, wird beispielsweise von Sirri und Tufano (1998), Jain und Wu (2000), Huber (2012) und McDonald und Rietz (2014) untersucht. Sirri und Tufano (1998) zeigen, dass Investoren, die in Aktienfonds investieren, sich an historischen Renditen orientieren. Aus diesem Grund werden gerade jene Fonds vermarktet, die in der Vergangenheit überproportionale Renditen erzielten (Jain und Wu (2000)). Del Guercio und Tkac (2008) finden heraus, dass amerikanische Fonds mit einer Top-Bewertung von Morningstar überproportionale Mittelzuflüsse verzeichnen. Bei sehr niedrig bewerteten Fonds sind stattdessen überproportionale Mittelabflüsse festzustellen. In Deutschland haben Veränderungen des Morningstar Ratings hingegen einen geringeren Einfluss auf das Investitionsverhalten. So zeigen Füss et al. (2010), dass Fonds bei Morningstar-Herabstufungen kaum Mittelabflüsse verzeichnen. Verbesserungen des Morningstar Ratings hin zu einer Top-Bewertung führen allerdings auch in Deutschland zu starken Mittelzuflüssen.

3 Aufbau und Funktionsweise der drei Fondsbewertungen

Gemeinsamkeiten und Unterschiede

Das Feri Trust Rating existiert seit Anfang 1999 und gehört zu den am weitesten verbreiteten Fondsbewertungen in Europa. Die Finanztest-Bewertung wird bereits seit 1991 von Stiftung Warentest, einer der bekanntesten Verbraucherorganisationen in Deutschland, veröffentlicht. Eine FondsNote erhalten Fonds hingegen erst seit Ende 2002.

Alle drei Fondsbewertungen verwenden fünf Bewertungsklassen, wobei sich die Bezeichnungen allerdings unterscheiden. Zum Beispiel erhalten die besten Fonds mit Feri Trust Rating ein (A), wohingegen sie bei der FondsNote mit der Note 1 versehen werden. Eine „stark überdurchschnittliche Bewertung“ erhält ein Top-Fonds mit Finanztest-Bewertung. Tabelle 1 stellt die fünf Bewertungsklassen der drei Fondsbewertungen gegenüber. Eine Top-Bewertung erhält ein Fonds nur dann, wenn er sich deutlich vom Durchschnitt seiner Peer Group absetzt. Dies ist ein bedeutender Unterschied zum Morningstar Rating, bei dem immer 10% der bewerteten Fonds eine Top-Bewertung erhalten.

Tabelle 1: Bewertungsklassen der drei Fondsbewertungen

Bewertungsklasse	Feri Trust Rating	Finanztest-Bewertung	FondsNote
Top-Bewertung	(A)	Stark überdurchschnittliche Bewertung	1
Zweithöchste Bewertung	(B)	Überdurchschnittliche Bewertung	2
Mittlere Bewertung	(C)	Durchschnittliche Bewertung	3
Zweitniedrigste Bewertung	(D)	Unterdurchschnittliche Bewertung	4
Niedrigste Bewertung	(E)	Stark unterdurchschnittliche Bewertung	5

Alle drei Fondsbewertungen unterteilen die zu untersuchenden Fonds entsprechend ihrem Anlageschwerpunkt in verschiedene Gruppen. Um einer Gruppe zugeordnet zu werden, muss ein Fonds mindestens 90% seines Fondsvermögens in diesen Anlageschwerpunkt investieren. Eine Zuordnung nach Investment Styles erfolgt im Vergleich zum Morningstar Rating nicht. Dies ist durchaus kritisch, da Fonds innerhalb einer Gruppe nur aufgrund ihres Investment Styles in der jeweiligen Marktphase eine Überrendite erzielen können.

Ein Vergleich der drei Fondsbewertungen in Tabelle 2 zeigt, dass sich die Bewertungen der Fonds innerhalb des Anlageschwerpunktes „Aktienfonds Deutschland“ voneinander unterscheiden können. Beispielsweise werden 29% beziehungsweise 19% der Fonds, die eine FondsNote erhalten, nicht mit einem Feri Trust Rating und einer Finanztest-Bewertung ausgezeichnet. Der Grund hierfür ist, dass sich der Bewertungszeitraum und das untersuchte Fondsuniversum zwischen den Fondsbewertungen unterscheiden. Fonds erhalten ein Feri Trust Rating und eine Finanztest-Bewertung, wenn sie seit mindestens fünf Jahren zum öffentlichen Vertrieb zugelassen

sind. Eine FondsNote gibt es bereits bei einer Historie von vier Jahren. Zu unterscheiden ist zudem nach Ländern, in denen ein Fonds zum Vertrieb zugelassen sein muss, damit dieser bewertet wird. Bei allen drei Fondsbewertungen werden die zu bewertenden Fonds in Deutschland vertrieben. Daneben werden beim Feri Trust Rating auch Fonds aus Österreich, Italien, Schweden, Frankreich, Großbritannien und der Schweiz mit einbezogen. Ein Feri Trust Rating erhalten zudem nur Fonds, die kürzlich keinen Strategiewechsel oder Managerwechsel vollzogen haben und eine Peer Group von mindestens 20 Fonds besitzen. Bei der FondsNote hingegen werden Fonds auch mit einbezogen, wenn Änderungen in der Strategie des Fonds und des Fondsmanagements vorliegen.

Tabelle 2: Einheitlichkeit der drei Fondsbewertungen

Panel A

		FondsNote					
		Q5	Q4	Q3	Q2	Q1	NR
Feri Trust	Q5	**0.30**	0.59	0.09	0.00	0.00	0.02
	Q4	0.07	**0.37**	0.37	0.05	0.03	0.11
	Q3	0.01	0.15	**0.46**	0.18	0.00	0.20
	Q2	0.00	0.03	0.21	**0.31**	0.01	0.43
	Q1	0.00	0.00	0.04	0.33	**0.20**	0.43
	NR	0.05	0.10	0.05	0.06	0.03	**0.71**

Panel B

		Finanztest					
		Q5	Q4	Q3	Q2	Q1	NR
Feri Trust	Q5	**0.44**	0.47	0.04	0.00	0.00	0.05
	Q4	0.03	**0.47**	0.42	0.01	0.00	0.07
	Q3	0.00	0.05	**0.59**	0.28	0.00	0.07
	Q2	0.00	0.00	0.18	**0.62**	0.12	0.08
	Q1	0.00	0.00	0.00	0.44	**0.55**	0.01
	NR	0.01	0.06	0.06	0.02	0.01	**0.84**

Panel C

		FondsNote					
		Q5	Q4	Q3	Q2	Q1	NR
Finanztest	Q5	**0.40**	0.50	0.06	0.00	0.00	0.04
	Q4	0.16	**0.49**	0.23	0.02	0.02	0.09
	Q3	0.03	0.21	**0.40**	0.14	0.01	0.22
	Q2	0.01	0.04	0.27	**0.31**	0.03	0.35
	Q1	0.00	0.00	0.07	0.43	**0.17**	0.33
	NR	0.01	0.05	0.05	0.06	0.02	**0.81**

Die drei Tabellen (Panel A-C) stellen die durchschnittlichen Anteile von Fonds in den jeweiligen Bewertungsklassen dar. Untersucht werden die Bewertungsklassen des Feri Trust Ratings, der Finanztest-Bewertung und der FondsNote. Q1 entspricht dem Anteil der Fonds mit der niedrigsten und Q5 mit der höchsten Bewertung. NR gibt den Anteil aller von einer Fondsbewertung nicht bewerteten Fonds wider. In Panel A werden die Fonds mit Feri Trust Rating mit denen der FondsNote verglichen. In Panel B werden die Fonds mit Feri Trust Rating der Finanztest-Bewertung gegenübergestellt. Panel C stellt den Vergleich der Fondsbewertungen der Finanztest-Bewertung und FondsNote dar. Der Untersuchungszeitraum umfasst sämtliche Fondsbewertungen von 12/2002 bis 12/2009.

Ein Fonds, der von allen drei Fondsbewertungen untersucht wird, erhält häufig unterschiedliche Bewertungen. Tabelle 2 zeigt, dass nur 30% beziehungsweise 44% der Fonds mit dem höchsten Feri Trust Rating auch die höchste FondsNote beziehungsweise Finanztest-Bewertung erhal-

ten. Dieser Zusammenhang besteht ebenso für niedriger bewertete Fonds. Da sich die Fondsbewertungen hinsichtlich eines Fonds unterscheiden, sollte auch die Vorhersagekraft voneinander abweichen. Ein Vergleich der drei Fondsbewertungen ist daher sinnvoll. Unter Umständen kann die Vorhersagekraft allerdings ähnlich sein, da einige Fonds nicht von allen drei Fondsbewertungen beurteilt werden. Der Grund hierfür ist, dass die Bewertung dieser Fonds einen Einfluss auf die Vorhersagekraft der Fondsbewertung hat.

Die Methodik des Feri Trust Ratings

Ist ein Fonds länger als fünf Jahre zum öffentlichen Vertrieb zugelassen, bewertet Feri Trust diesen Fonds rein quantitativ mithilfe von Performance- und Risikokennzahlen. Beträgt die Historie weniger als fünf Jahre, wird ein Fonds nur bewertet, wenn der Emittent des Fonds einen Auftrag an Feri Trust erteilt. Trifft dies zu, dann werden neben quantitativen auch qualitative Faktoren berücksichtigt. Je kürzer der historische Zeitraum eines Fonds, desto größer ist das Gewicht qualitativer Faktoren (Feri Trust (2002)). Die Bewertung mit qualitativen Faktoren spielt für den in dieser Arbeit untersuchten Datensatz allerdings keine Rolle.

Für jeden Fonds innerhalb einer Gruppe werden insgesamt 12 Kennzahlen, jeweils sechs Performance- und Risikokennzahlen, ermittelt (Tabelle 3). Die sechs Performancekennzahlen, die zu 70% in den Bewertungsprozess einfließen, setzen sich aus Faktoren der relativen Performance, der langfristigen Ertragskraft und der Stabilität zusammen. Die übrigen 30% bestehen aus sechs Risikokennzahlen wie dem Timingrisiko, dem Verlustrisiko und dem Verhaltensrisiko. Jede Kennzahl eines Fonds wird normiert und somit einer Punkteskala von 1 bis 100 Punkten zugeordnet. Der Median-Fonds erhält bezogen auf eine Kennzahl immer 50 Punkte. Das Gesamturteil eines Fonds ergibt sich abschließend aus dem gewichteten Mittel aller zwölf Kennzahlen, die in Tabelle 3 dargestellt sind.

Tabelle 3: Bewertungsmethodik des Feri Trust Ratings

Performancekennzahlen	**70%**	**Risikokennzahlen**	**30%**
Relative Performance		*Timingrisiko*	
Outperformance geg. Index p.a.	15%	Volatilität p.a.	25%
Durchschn. Rang in rollierenden Drei-Monats-Perioden	20%	*Verlustrisiko*	
Langfristige Ertragskraft		Max. gleitender Verlust in 6 Monaten	20%
Positive Elastizität	15%	Wahrscheinlichkeit eines Verlustmonats	10%
Differenz der positiven und negativen Elastizität	10%	Durchschn. Verlust der Verlustmonate	10%
Stabilität		Negative Elastizität	20%
Wahrscheinlichkeit der Outperformance geg. Index	20%	*Verhaltensrisiko*	
Wahrscheinlichkeit der Outperformance geg. Peer Group	20%	Tracking Error	15%

Eine Top-Bewertung erhalten Fonds, die mindestens 78 von 100 Punkten erzielen. Dies kann zur Folge haben, dass kein einziger Fonds innerhalb einer Gruppe ein (A)-Rating erhält. Ein Fonds

erhält ein (B)-Rating, wenn er mindestens 60 Punkte erreicht, ein (C)-Rating für mindestens 41 Punkte, ein (D)-Rating ab 23 Punkten und das niedrigste Rating für weniger als 23 Punkte.

Die Methodik der Finanztest-Bewertung

Die Bewertung von Fonds eines Anlageschwerpunktes erfolgt bei der Finanztest-Bewertung ausschließlich mithilfe von zwei quantitativen Faktoren $R_{i,t}$ und $SR_{i,t}$. $R_{i,t}$ ist die Summe aller positiven Renditedifferenzen zwischen Fonds *i* und seiner Peer Group zum Zeitpunkt *t* in Relation zur Summe aller absoluten Renditedifferenzen. $SR_{i,t}$ entspricht dem arithmetischen Mittel von $R_{i,t-48}$ bis $R_{i,t}$.

$R_{i,t}$ wird ermittelt als:

$$R_{i,t} = 100 \cdot \frac{\sum_{\tau=t-60}^{t-1} D_{i,\tau} \cdot \left(r_{i,\tau} - r_{Peer\ group,\tau}\right)}{\sum_{\tau=t-60}^{t-1} \left|r_{i,\tau} - r_{Peer\ group,\tau}\right|} \tag{1}$$

und misst die relative Wertentwicklung von Fonds *i* im Vergleich zu seiner Peer Group in den zurückliegenden 60 Monaten. Hierbei ist $r_{i,\tau}$ die monatliche Rendite von Fonds *i* und $r_{Peer\ group,\tau}$ die durchschnittliche monatliche Rendite der Peer Group jeweils zum Zeitpunkt τ. $D_{i,\tau}$ entspricht einer Dummy-Variable, die nur dann den Wert 1 annimmt, wenn die Fondsrendite größer als die Rendite der Peer Group ist. Ansonsten ist die Dummy-Variable gleich Null.

Für $SR_{i,t}$ folgt hieraus:

$$SR_{i,t} = \frac{1}{49} \cdot \sum_{\tau=t-48}^{t} R_{i,\tau}. \tag{2}$$

Hierbei wird unterstellt, dass der 60 monatige Untersuchungszeitraum in 49 monatlich rollierende Ein-Jahres-Zeiträume zerlegt werden kann. Nur diese Zeiträume von t bis $t - 48$ fließen in die Berechnung mit ein.

Die Gesamtbewertung $Pkt\_FB_{i,t}$ von Fonds *i* zum Zeitpunkt *t* bildet sich aus dem gewichteten Mittel der beiden Variablen $R_{i,t}$ und $SR_{i,t}$. $R_{i,t}$ fließt zu 75% und $SR_{i,t}$ zu 25% in die Gesamtbewertung ein.

$$Pkt\_FB_{i,t} = 0{,}75 \cdot R_{i,t} + 0{,}25 \cdot SR_{i,t}, \tag{3}$$

Die Finanztest-Bewertung $Pkt\_FB_{i,t}$ liegt zwischen 0 und 100 Punkten. Ein Fonds wird als stark überdurchschnittlich bezeichnet, wenn er mindestens 65 Punkte erzielt. Eine überdurchschnittliche Bewertung wird zwischen 55 und 64,99 Punkte erreicht. Durchschnittliche und unterdurchschnittliche Bewertungen erhalten jene Fonds, die 45 bis 54,99 Punkte beziehungsweise 35 bis 44,99 Punkte erzielen. Fonds mit weniger als 35 Punkten erhalten eine stark unterdurchschnittliche Bewertung.

Die Methodik der FondsNote

Die FondsNote kann im Vergleich zum Feri Trust Rating und der Finanztest-Bewertung sowohl quantitative als auch qualitative Faktoren enthalten. Zunächst erfolgt eine Basisbewertung, die ausschließlich quantitative Faktoren beinhaltet. Dabei werden alle Fonds eines Anlageschwerpunktes mit der relativen Performance gegenüber dem Index, der relativen Performance gegenüber der Peer Group und der Standardabweichung der monatlichen Renditen analysiert. Erhält ein Fonds mindestens die dritthöchste Basisbewertung, erfolgt anschließend eine Bewertung mit qualitativen Faktoren wie der Managerkontinuität, der Konstanz des Investmentansatzes und der Produktwahrheit/Produktklarheit (Axel Springer und FondsConsult (2007)).

Die Basisbewertung basiert auf einem Untersuchungszeitraum von 48 Monaten. Dieser wird in jeweils 36 monatlich rollierende Ein-Jahres-Zeiträume zerlegt. Für jeden dieser Ein-Jahres-Zeiträume werden die drei quantitativen Faktoren bestimmt und in Abhängigkeit der Peer Group einer Punkteskala von 0 bis 100 Punkten zugeordnet. Ein Fonds erhält beispielsweise die Höchstpunktzahl bezüglich der relativen Performance, wenn er die höchste Überrendite aufweist. Die Punktzahl von Fonds *i* bezüglich des quantitativen Faktors *j* zum Zeitpunkt t ist mit $Pkt_{i,j,t}$ definiert. Daraus folgt, dass:

$$Faktor_{i,j,t} = \sum_{\tau=t-36}^{t-1} Pkt_{i,j,\tau} \cdot [1 - 0{,}05 \cdot (t - \tau)], \tag{4}$$

wobei $Pkt_{i,j,\tau}$ für näherliegende Zeitpunkte ein höheres Gewicht erhält als für weiter zurückliegende. Die Gesamtpunktzahl der Basisbewertung $Pkt\_FN_{i,t}$ für einen Fonds *i* zum Zeitpunkt t entspricht letztlich dem gewichteten Mittel der drei quantitativen Faktoren ($Faktor_{i,j,t}$). Das Gewicht G_j für den Faktor der relativen Performance gegenüber dem Index beträgt 45%. Der Faktor der relativen Performance gegenüber der Peer Group und der Faktor der Standardabweichung der Renditen beträgt 30% beziehungsweise 25%. Daraus folgt, dass:

$$Pkt\_FN_{i,t} = \sum_{j=1}^{3} Faktor_{i,j,t} \cdot G_j. \tag{5}$$

Aus $Pkt\_FN_{i,t}$ werden die Basisbewertungen mit Noten von 1 bis 5 abgeleitet. Basisbewertungen von Fonds eines Anlageschwerpunktes sind normalverteilt. Die Note 1 entspricht der höchsten Basisbewertung, die Note 5 der Niedrigsten. Für Fonds mit Noten von 4 bis 5 ist die Basisbewertung gleichzeitig die Endbewertung.

Fonds erhalten eine qualitative Bewertung, wenn sie zum Bewertungszeitpunkt mindestens die dritthöchste Basisbewertung erzielen. Qualitative Faktoren können allerdings nicht zu einer Verbesserung, sondern nur zu einer Verschlechterung der Endbewertung führen. Die qualitativen Faktoren fließen zu 30% in die Endbewertung mit ein. Das Gewicht der Basisbewertung reduziert sich somit auf 70%.

Tabelle 4 stellt sämtliche qualitativen Faktoren und deren Gewichtung vor. Ein Fonds erhält die Höchstnote bezüglich der Managerkontinuität und der Konstanz des Investmentansatzes, wenn sich beide Faktoren seit mindestens vier Jahren nicht verändert haben. Die Beurteilung der Produktwahrheit/Produktklarheit untergliedert sich in vier Subkriterien, für die ein Fonds jeweils einen Punkt erhält, sollte das Kriterium erfüllt sein.

Tabelle 4: Bewertungsmethodik der FondsNote – Qualitative Faktoren

Qualitativer Faktor	Gewicht
Managerkontinuität	40%
Konstanz des Investmentansatzes	40%
Produktwahrheit/Produktklarheit	20%
- Fondsname aussagekräftig?	
- Anlagerichtlinien konstant?	
- Index geeignet?	
- Depotstruktur gibt Fondsnamen wider?	

4 Daten und Methodik

4.1 Datengrundlage

Der Analyse liegen die Fondsbewertungen des Feri Trust Ratings, der Finanztest-Bewertung und der FondsNote zu Aktienfonds mit Anlageschwerpunkt Deutschland zugrunde. Die Fondsbewertungen aller Fonds, die zum Vertrieb zugelassen waren, stehen für den Zeitraum von Dezember 2002 bis Dezember 2009 vollständig zur Verfügung und stammen direkt von den Anbietern der Fondsbewertungen. Ein Survivorship-Bias, der bei fehlenden Fondsbewertungen entstehen würde, liegt somit nicht vor. Fondspreise sowie Fondsvermögen sind aus Thomson Reuters

Datastream entnommen. Weitere Daten wie das Total Expense Ratio (TER) wurden vom BVI und der Zeitschrift Fondsxpress, die auch die FondsNote veröffentlicht, zur Verfügung gestellt. Für die Ermittlung der Performancemaße werden die von Brückner et al. (2014) für Deutschland ermittelten Fama/French Faktoren verwendet.

Der Datensatz besteht aus insgesamt 105 Fonds mit Feri Trust Rating, 139 Fonds mit Finanztest-Bewertung und 96 Fonds mit FondsNote. Die Anzahl der jeweils untersuchten Fonds für verschiedene Jahre innerhalb des Untersuchungszeitraumes ist in Tabelle 5 dargestellt. Hierbei fällt auf, dass sich die Anzahl der bewerteten Fonds mit Feri Trust Rating und Finanztest-Bewertung im Zeitablauf verringert. Vor allem niedrig bewertete Fonds fusionieren oder werden liquidiert. Als Gründe hierfür sind sowohl Kosteneinsparungen und Effizienzsteigerungen auf Emittentenebene als auch eine Steigerung der Reputation des Emittenten anzuführen, indem beispielsweise der Vertrieb von Fonds mit geringer Performance während der Finanzkrise eingestellt wurde. Die Anzahl der Fonds mit FondsNote erhöht sich allerdings im Vergleich zu den beiden anderen Bewertungen. Als ein Grund kann der stetige Aufbau des Universums der zu bewerteten Fonds ab September 2002, dem Zeitpunkt der erstmaligen Veröffentlichung der FondsNote, genannt werden.

Tabelle 5: Datengrundlage

	Feri Trust	Finanztest	FondsNote
Zeitpunkt			
12/2002	81	89	58
12/2003	75	80	65
12/2004	71	76	75
12/2005	70	80	78
12/2006	68	81	78
12/2007	69	80	80
12/2008	66	85	92
12/2009	59	75	90

Die Tabelle stellt die Anzahl untersuchter Aktienfonds mit Anlageschwerpunkt Deutschland zum Jahresende innerhalb des Untersuchungszeitraumes dar.

Fonds, die liquidiert und fusioniert wurden, können bei der Bewertung der Vorhersagekraft der drei Fondsbewertungen zu Verzerrungen führen. Der Grund hierfür ist, dass vor allem niedriger bewertete Fonds aufgelöst werden oder mit anderen Fonds fusionieren. Der hierbei entstehende Survivorship Bias wurde in der Literatur auf verschiedene Weise reduziert (Vgl. Blake und Morey (2000), Morey und Gottesman (2006) und Kräussl und Sandelowski (2007)). In Anlehnung an die Methodik von Morey und Gottesman (2006) werden liquidierte Fonds sowie Fonds, die ihren Anlageschwerpunkt nach der Bewertung ändern, mit Renditen sehr ähnlicher Fonds (identische

Fondsbewertung und vergleichbare einjährige risikoadjustierte Rendite) fortgeführt. Fonds, die mit anderen Fonds fusionieren, werden mit Renditen des weiter bestehenden Fonds fortgesetzt.

4.2 Methodik

Dummy-Variablen Regression

Die Vorhersagekraft einer Fondsbewertung lässt sich anhand der durchschnittlichen zukünftigen Performance der Bewertungsklassen messen. Ebenso wie bei Blake und Morey (2000), Morey und Gottesman (2006) und Füss et al. (2010) werden zunächst die Fonds anhand ihrer Fondsbewertung in fünf Bewertungsklassen eingeteilt. In dieser Arbeit wird hierfür der Dezember eines jeden Jahres verwendet. Auf dieser Basis wird für jeden Fonds die zukünftige Performance für unterschiedliche Zeiträume zwischen Januar 2003 und Dezember 2010 ermittelt. Die untersuchten Performancezeiträume umfassen somit insgesamt acht sich nicht überlappende Ein-Jahres-Zeiträume. Um die Stabilität der Performance und die Vorhersagekraft der Fondsbewertungen für längerfristige Zeiträume zu untersuchen, werden auch die sich zeitlich überlappenden sechs Drei-Jahres- und vier Fünf-Jahres-Zeiträume einbezogen. Dies ist möglich, da die Zusammensetzung der Bewertungsklassen mit Fonds von Jahr zu Jahr stark variiert.

Um die zukünftige Performance der Bewertungsklassen zu analysieren, werden Querschnittsregressionen mit Dummy-Variablen durchgeführt. Hierbei wird die durchschnittliche Performance der Bewertungsklassen miteinander verglichen. Die Dummy-Variablen Regression besitzt die Form:

$$S_i = \beta_1 + \beta_2 D2_i + \beta_3 D3_i + \beta_4 D4_i + \beta_5 D5_i + \varepsilon_i, \tag{6}$$

wobei S_i der Performance von Fonds *i* für den untersuchten Zeitraum entspricht. Die Performance wird mithilfe des annualisierten 4-Faktor-Alphas und des annualisierten geometrischen Mittels (basierend auf monatlichen Renditen) gemessen. $D2_i$ bis $D5_i$ sind Dummy-Variablen, die die Bewertungsklasse von Fonds *i* wiederspiegeln. Zum Beispiel ist $D5_i$ für Fonds *i* gleich 1, wenn dieser Fonds die höchste Fondsbewertung erhalten hat. $D5_i$ ist 0, wenn dies nicht zutrifft. β_1 bildet die Referenzklasse, wenn sämtliche Dummy-Variablen gleich 0 sind und gibt die durchschnittliche Performance der am niedrigsten bewerteten Fonds wider. Folglich wird die Performance der Bewertungsklassen im Verhältnis zur Referenzklasse gemessen. Letztlich entspricht ε_i dem Störterm der Regression.

Eine niedrige Fondsbewertung drückt aus, dass dieser Fonds zukünftig eine geringere Performance aufweisen sollte als besser bewertete Fonds. Demnach besitzt eine Fondsbewertung eine

Vorhersagekraft, wenn β_2 bis β_5 signifikant und positiv sind. Zudem sollten die Koeffizienten die Ungleichung $\beta_2 < \beta_3 < \beta_4 < \beta_5$ erfüllen. Im Vergleich zur Literatur, die hauptsächlich das Morningstar Rating untersucht, werden in dieser Arbeit nicht die Fonds mit der höchsten Fondsbewertung als Referenzklasse verwendet. Der Grund hierfür ist, dass die drei Fondsbewertungen nicht immer mindestens einen Fonds dieser Bewertungsklasse zuordnen.

Alternative Fondsbewertung mittels Performancemaßen

Als nächstes soll die Vorhersagekraft der drei Fondsbewertungen mit der Vorhersagekraft zweier Performancemaße, des 4-Faktor-Alphas und des geometrischen Mittels der Renditen, verglichen werden. Sollten die beiden Performancemaße eine höhere Vorhersagekraft als die Fondsbewertungen aufweisen, stellt sich die Frage nach der Existenzberechtigung der Fondsbewertungen. Umgekehrt kann gelten, dass Fondsbewertungen einen zusätzlichen Informationsgehalt für Investoren liefern, wenn ihre Vorhersagekraft höher als die der Performancemaße ist.

Des Weiteren kann hierdurch untersucht werden, ob sich die Performance eines Fonds über die Zeit hinweg ändert oder nahezu gleich bleibt. Unterliegt die Performance eines Fonds ständigen Änderungen, dann haben Fondsbewertungen wie das Feri Trust Rating und die Finanztest-Bewertung eine geringe Vorhersagekraft. Der Grund hierfür ist, dass diese beide Fondsbewertungen ausschließlich auf historischen Rendite- und Risikokennzahlen basieren. Liegt keine Beständigkeit in der Performance der untersuchten Fonds vor, könnte die FondsNote dennoch eine Vorhersagekraft besitzen, da sie neben quantitativen auch qualitative Faktoren mit einbezieht.

Zunächst wird die Performance der zu bewertenden Fonds für einen festgelegten Zeitraum bestimmt. Um die Vergleichbarkeit zwischen den Fondsbewertungen zu gewährleisten, wird einheitlich ein Fünf-Jahres-Zeitraum verwendet, da das Feri Trust Rating und die Finanztest-Bewertung auch auf diesem Zeitraum basieren. Alternativ wurden auch Ein- und Drei-Jahres-Zeiträume untersucht, die allerdings zu einer geringeren Vorhersagekraft der Performancemaße führten. Des Weiteren werden nur Fonds mit einbezogen, die zum Bewertungszeitpunkt von jeweils einer der drei Fondsbewertungen bewertet wurden.

Um einen Vergleich mit den drei Fondsbewertungen durchführen zu können, werden die Fonds basierend auf ihrer Performance in fünf Bewertungsklassen eingeteilt. Dabei entspricht die Größe einer Bewertungsklasse dem durchschnittlichen Anteil von Fonds, die jeweils von einer der drei Fondsbewertungen innerhalb des Untersuchungszeitraumes (12/2002-12/2009) dieser Bewertungsklasse zugeordnet wurden. Zum Beispiel erhielten durchschnittlich 14% aller von Feri Trust bewerteten Fonds eine Top-Bewertung. Dies gilt ebenso für Fonds mit der höchsten Finanztest-Bewertung, wohingegen nur 5% aller Fonds die höchste FondsNote erhielten.

Mithilfe von Spearmans Rangkorrelationskoeffizient werden nun die neu gebildeten Bewertungsklassen mit den drei Fondsbewertungen verglichen. Der Rangkorrelationskoeffizient gibt darüber Auskunft, wie stark die neuen Bewertungsklassen mit den Fondsbewertungen korreliert sind. Daraus lässt sich ableiten, ob sich die Vorhersagekraft der Performancemaße überhaupt von den drei Fondsbewertungen unterscheidet. Wenn das Feri Trust Rating, die Finanztest-Bewertung und die FondsNote individuelle Fondsbewertungen sind, dann sollte die Korrelation zu den Performancemaßen gering sein. Abschließend werden ebenso wie für die drei Fondsbewertungen Regressionen mit Dummy-Variablen durchgeführt, um die Vorhersagekraft der beiden Performancemaße zu ermitteln. Dabei wird die zukünftige Performance der Fonds mithilfe des gleichen Performancemaßes bestimmt, das auch für die Bildung der neuen Bewertungsklassen verwendet wurde. Die Vorhersagekraft von Bewertungsklassen, die beispielsweise auf dem 4-Faktor-Alpha basieren und deren zukünftige Performance mit dem geometrischen Mittel bestimmt wird, ist zudem untersucht worden. Allerdings werden diese Ergebnisse nicht vorgestellt, da die Vorhersagekraft geringer ist.

Einfluss von Fondsvermögen und Total Expense Ratio

Als nächstes soll untersucht werden, welche Beziehung zwischen einer Fondsbewertung und dem Fondsvermögen besteht. Zum Beispiel sollten Fonds mit Top-Bewertung im Vergleich zu niedriger bewerteten Fonds überproportionale Mittelzuflüsse verzeichnen. Sehr niedrig bewertete Fonds sollten hingegen Mittelabflüsse aufweisen. Hierdurch lässt sich auch untersuchen, ob Veränderungen des Fondsvermögens die Vorhersagekraft der drei Fondsbewertungen beeinflussen. Wird beispielsweise ein Fonds mit einer Top-Bewertung ausgezeichnet, dann sollte dieser Fonds starke Mittelzuflüsse verzeichnen. Diese Mittel könnten allerdings aufgrund von Investitionsbeschränkungen und mangelnder rentierlicher Alternativen weniger stark gewinnbringend investiert werden, wodurch sich die Performance des Fonds und die Vorhersagekraft der Fondsbewertung verringern würde. Eine Voraussetzung hierfür ist, dass Investoren Fondsbewertungen in ihren Entscheidungsprozess beim Kauf oder Verkauf von Fondsanteilen mit einbeziehen.

Das Feri Trust Rating sollte den stärksten Einfluss auf das Investitionsverhalten haben und somit auch das Fondsvermögen beeinflussen. Zu ähnlichen Ergebnissen sollte die Finanztest-Bewertung führen, wohingegen bei der FondsNote der geringste Einfluss zu erwarten ist. Der Grund für diese Reihenfolge ist, dass das Feri Trust Rating neben dem Morningstar Rating zu den bekanntesten Fondsbewertungen in Europa zählt. Die Finanztest-Bewertung gilt aufgrund ihrer öffentlichen Wahrnehmung in Deutschland als ein weiterer großer Anbieter von Fondsbewertun-

gen. Die FondsNote ist im Vergleich zu den beiden anderen Fondsbewertungen weniger stark bekannt.

Das Investitionsverhalten wird in der Literatur am häufigsten mithilfe des net fund flows eines Fonds untersucht (Vgl. z. Bsp. Del Guercio und Tkac (2008) und Füss et al. (2010)). Um den Einfluss der Fondsbewertungen zu untersuchen, werden für jeden Fonds einer Bewertungsklasse die monatlichen Veränderungen im Fondsvermögen ermittelt. Hierbei werden im Vergleich zu Füss et al. (2010) Veränderungen, die nicht auf den Kauf und Verkauf von Fondsanteilen zurückgeführt werden können, nicht berücksichtigt. Für die Veränderungen im Fondsvermögen $Flow_{i,t}$ von Fonds *i* zum Zeitpunkt *t* gilt:

$$Flow_{i,t} = TNA_{i,t} - TNA_{i,t-1} \cdot (1 + R_{i,t}), \tag{7}$$

wobei $TNA_{i,t}$ dem Fondsvermögen und $R_{i,t}$ der monatlichen Rendite von Fonds *i* zum Zeitpunkt *t* entspricht. Fondsbewertungen haben keinen Einfluss auf die durchschnittliche Veränderung des Fondsvermögens einer Bewertungsklasse, wenn sich die Bewertungsklassen hinsichtlich dieser Kennzahl nicht voneinander unterscheiden. Demnach sollte die nachfolgende Gleichung erfüllt sein:

$$\overline{Flow_{Q1}} = \overline{Flow_{Q2}} = \overline{Flow_{Q3}} = \overline{Flow_{Q4}} = \overline{Flow_{Q5}}, \tag{8}$$

wobei $\overline{Flow_{Q1}}$ die durchschnittliche monatliche Veränderung des Fondsvermögens der am niedrigsten bewerteten Fonds bezeichnet. $\overline{Flow_{Q2}}$, $\overline{Flow_{Q3}}$, $\overline{Flow_{Q4}}$ und $\overline{Flow_{Q5}}$ spiegeln die Veränderungen des Fondsvermögens der am zweitniedrigsten bis am höchsten bewerteten Fonds wider. Ähnlich wie bei Füss et al. (2010) sollen neben dem Gesamtzeitraum auch Marktphasen mit steigenden und sinkenden DAX-Kursen untersucht werden (Bullenmarkt: 12/2002-06/2007, Bärenmarkt: 12/2007-02/2009), um mögliche Veränderungen im Investitionsverhalten zu analysieren.

Neben der Analyse des Einflusses von Fondsbewertungen auf das Investitionsverhalten soll auch der Einfluss der Fondskosten auf die Vorhersagekraft der drei Fondsbewertungen untersucht werden. Beispielsweise zeigen Blake und Morey (2000) und Füss et al. (2010), dass insbesondere die am niedrigsten bewerteten Fonds eine geringere Performance als höher bewertete Fonds aufweisen. Was zunächst auf eine Vorhersagekraft der Fondsbewertung hindeutet, könnte allerdings alleine durch die Fondskosten erklärt werden, wenn die am niedrigsten bewerteten Fonds

die höchsten Kosten aufweisen. Des Weiteren hat die Literatur gezeigt, dass sich die Performance von Fonds mit Top-Bewertung kaum von jener mit zweithöchster Bewertung unterscheidet. Zu vermuten ist, dass die am höchsten bewerteten Fonds höhere Kosten als die am zweithöchsten bewerteten Fonds haben.

Wenn die Fondskosten nicht der Grund dafür sind, weshalb sich die Performance zwischen den Bewertungsklassen kaum unterscheidet, dann sollten sich auch die durchschnittlichen Fondskosten zwischen den Bewertungsklassen nicht voneinander unterscheiden.

$$\overline{TER_{Q1}} = \overline{TER_{Q2}} = \overline{TER_{Q3}} = \overline{TER_{Q4}} = \overline{TER_{Q5}}, \tag{9}$$

wobei $\overline{TER_{Q1}}$ dem durchschnittlichen Total Expense Ratio (TER) aller Fonds mit der niedrigsten Bewertung entspricht. $\overline{TER_{Q2}}$, $\overline{TER_{Q3}}$, $\overline{TER_{Q4}}$, $\overline{TER_{Q5}}$ spiegeln aufsteigend die Fondskosten der anderen vier Bewertungsklassen wider. Der Vergleich der Fondskosten erfolgt im Dezember eines jeden Jahres von 2005 bis 2009. Der Grund für den eingeschränkten Zeitraum ist, dass die Veröffentlichung von Fondskosten erst ab 2003 gesetzlich vorgeschrieben ist. Zudem sind für die Jahre vor 2005 erhebliche Datenlücken vorzufinden, die die Aussagekraft einschränken würden. Die Total Expense Ratios liegen für 77% bis 96% aller Fonds innerhalb des Untersuchungszeitraumes vor.

5 Ergebnisse

5.1 Prognosefähigkeit der drei Fondsbewertungen

Die Tabellen A1.1, A1.2 und A1.3 zeigen die Ergebnisse der Regression mit Dummy-Variablen für das Feri Trust Rating, die Finanztest-Bewertung und die FondsNote. Als Performancemaß wurde das annualisierte 4-Faktor-Alpha und das geometrische Mittel verwendet. Zu beachten ist, dass bei der Finanztest-Bewertung nicht zu jedem Zeitpunkt Fonds mit einer Top-Bewertung vorhanden sind. In diesem Fall wird die Performance der Fonds auf drei Dummies regressiert.

Die Ergebnisse zeigen, dass das Feri Trust Rating und die Finanztest-Bewertung eine geringe Vorhersagekraft besitzen. Die F-Tests sind oftmals insignifikant. Zwischen den Bewertungsklassen liegen somit kaum Performanceunterschiede vor. Die Reihung der Koeffizienten, so dass $\beta_2 < \beta_3 < \beta_4 < \beta_5$, ist nur selten erfüllt. Dies gilt für beide Performancemaße, basierend sowohl auf Ein- und Drei-Jahres-Zeiträumen als auch auf Fünf-Jahres-Zeiträumen. Blake und Morey (2000) und Füss et al. (2010) gelangen für das Morningstar Rating zu ähnlichen Ergebnissen, ebenso wie

Müller und Weber (2014) für die Finanztest-Bewertung mit Anlageschwerpunkt „Aktienfonds Deutschland“.

Die FondsNote besitzt im Vergleich zum Feri Trust Rating als auch der Finanztest-Bewertung eine deutlich höhere Vorhersagekraft. Die F-Tests sind mit dem 4-Faktor-Alpha und dem geometrischen Mittel in 10 beziehungsweise 11 von 18 Untersuchungszeiträumen signifikant. Insbesondere bei den Drei- als auch Fünf-Jahres-Zeiträumen sind zahlreiche Performanceunterschiede zwischen den einzelnen Bewertungsklassen zu verzeichnen. Obwohl sich diese Zeiträume überlappen, kann hier auf eine höhere Vorhersagekraft der FondsNote geschlossen werden. Der Grund hierfür ist, dass ein mit einer FondsNote bewerteter Fonds seine Bewertungsklasse kurz- bis mittelfristig ändert. Meinhardt (2009) zeigt, dass ein Fonds mit Top-Bewertung bereits nach einem Jahr mit einer Wahrscheinlichkeit von 50% nicht mehr die höchste Bewertung erhält. Die untersuchten Fonds in den jeweiligen Bewertungsklassen unterscheiden sich daher stark zwischen den einzelnen Untersuchungszeiträumen.

Die Ergebnisse der FondsNote zeigen, dass insbesondere die am niedrigsten bewerteten Fonds (β_1) häufig die geringste Performance im Vergleich zu höher bewerteten Klassen aufweisen. Zudem ist die Ungleichung $\beta_2 < \beta_3 < \beta_4 < \beta_5$ häufiger erfüllt, so dass beispielsweise Fonds mit Top-Bewertung eine höhere Performance als die am zweithöchsten bewerteten Fonds haben. Eine Aussage über einen statistisch signifikanten Performanceunterschied kann zwischen diesen beiden Bewertungsklassen allerdings nicht getroffen werden, da bei der Dummy-Variablen Regression nur Vergleiche zur Referenzklasse (β_1) durchgeführt werden.

Ein Grund für die deutlich höhere Vorhersagekraft der FondsNote gegenüber den anderen beiden Fondsbewertungen könnte auf die Bewertungsmethodik zurückgeführt werden. Sowohl das Feri Trust Rating als auch die Finanztest-Bewertung beruhen ausschließlich auf quantitativen Faktoren, die rein vergangenheitsorientiert sind. Im Vergleich hierzu bezieht die FondsNote zusätzlich qualitative Faktoren mit ein, die eine Prognose über die zukünftige Performanceentwicklung eines Fonds ermöglichen.

5.2 Prognosefähigkeit von Performancemaßen

In Tabelle 6 sind die Spearmans Rangkorrelationskoeffizienten dargestellt. Diese zeigen, ob die Fondsbewertungen des Feri Trust Ratings, der Finanztest-Bewertung und der FondsNote mit den Bewertungsklassen, basierend auf dem 4-Faktor-Alpha und dem geometrischen Mittel, übereinstimmen.

Die Ergebnisse zeigen, dass die drei Fondsbewertungen für zahlreiche Zeitpunkte mit den neuen Bewertungsklassen stark positiv korreliert sind. Der stärkste Zusammenhang besteht, wenn die

Bewertungsklassen mithilfe des geometrischen Mittels gebildet wurden. Interessanterweise nehmen die Korrelationen für spätere Zeitpunkte ab, werden oftmals sogar insignifikant. Daraus lässt sich ableiten, dass die Fondsbewertungen nicht vollständig identisch sind. Eine Analyse der Vorhersagekraft von Fondsbewertungen, die auf beiden Performancemaßen basiert, sollte daher zu Veränderungen in der Vorhersagekraft führen.

Tabelle 6: Spearmans Rangkorrelationskoeffizient

	4-Faktor-Alpha			Geometrisches Mittel		
	Feri Trust	**Finanztest**	**FondsNote**	**Feri Trust**	**Finanztest**	**FondsNote**
12/2002	**0.766*** **	**0.639*** **	**0.712*** **	**0.788*** **	**0.848*** **	**0.759*** **
	(9.146)	(7.002)	(7.305)	(9.846)	(13.503)	(8.717)
12/2003	**0.725*** **	**0.780*** **	**0.617*** **	**0.745*** **	**0.888*** **	**0.716*** **
	(8.090)	(10.519)	(5.648)	(8.570)	(16.297)	(7.672)
12/2004	**0.862*** **	**0.795*** **	**0.543*** **	**0.900*** **	**0.899*** **	**0.778*** **
	(13.062)	(11.046)	(4.662)	(15.819)	(17.298)	(9.255)
12/2005	**0.721*** **	**0.784*** **	**0.337** **	**0.872*** **	**0.891*** **	**0.615*** **
	(8.000)	(10.656)	(2.578)	(13.715)	(16.571)	(5.838)
12/2006	**0.518*** **	**0.349*** **	0.163	**0.871*** **	**0.662*** **	**0.500*** **
	(4.278)	(2.840)	(1.193)	(12.545)	(6.732)	(4.325)
12/2007	**0.424*** **	0.111	**0.260* **	**0.880*** **	0.196	**0.565*** **
	(3.312)	(0.847)	(1.938)	(13.107)	(1.522)	(5.123)
12/2008	0.085	0.003	**0.289** **	**0.788*** **	0.042	**0.299** **
	(0.601)	(0.021)	(2.180)	(9.050)	(0.317)	(2.345)
12/2009	0.119	**0.301** **	-0.196	**0.595*** **	**0.571*** **	**0.247* **
	(0.851)	(2.400)	(-1.441)	(5.241)	(5.295)	(1.911)

Die Tabelle zeigt die Rangkorrelationskoeffizienten zwischen den neu gebildeten Bewertungsklassen, die auf dem 4-Faktor-Alpha und dem geometrischen Mittel basieren, und den drei Fondsbewertungen (Feri Trust Rating, Finanztest-Bewertung und FondsNote). Die beiden Performancemaße wurden für Fünf-Jahres-Zeiträume bestimmt. Die Nullhypothese besagt, dass zwischen den neu gebildeten Bewertungsklassen und den drei Fondsbewertungen kein Zusammenhang (keine Korrelation) besteht. *, ** und *** zeigen die Signifikanz auf dem 10%, 5% und 1% Niveau an.

Die Tabellen A2.1, A2.2 und A2.3 zeigen die Ergebnisse der neu gebildeten Bewertungsklassen, die auf dem 4-Faktor-Alpha und dem geometrisches Mittel basieren. Den Analysen liegt wiederum eine Regression mit Dummy-Variablen für Ein-, Drei- und Fünf-Jahres-Zeiträume zugrunde. Um einen Vergleich der Vorhersagekraft mit den drei Fondsbewertungen zu ermöglichen, werden nur die Fonds bewertet, die auch jeweils zum Bewertungszeitpunkt ein Feri Trust Rating, eine Finanztest-Bewertung oder eine FondsNote erhielten.

Die Bewertung mit beiden Performancemaßen führt im Vergleich zum Feri Trust Rating kaum zu einer Verbesserung der Vorhersagekraft. Die F-Tests der Dummy-Variablen Regression sind häufig insignifikant. Nur einzelne Bewertungsklassen erreichen eine signifikant höhere Performance als die Referenzklasse, die die am niedrigsten bewerteten Fonds abbildet. Im Vergleich zur Finanztest-Bewertung führen die Performancemaße zu einer stärkeren Differenzierung der Be-

wertungsklassen. Vor allem Fonds, die mithilfe des geometrischen Mittels in Bewertungsklassen eingeteilt wurden, unterscheiden sich hinsichtlich ihrer Performance deutlich stärker.

im Vergleich zu den Ergebnissen der FondsNote verbessert sich die Vorhersagekraft der Performancemaße nicht. Vielmehr ist zu beobachten, dass insbesondere die Einteilung von Fonds mit dem 4-Faktor-Alpha zu einer geringeren Differenzierung zwischen den Bewertungsklassen führt. Der Grund hierfür ist, dass das 4-Faktor-Alpha ebenso wie das geometrische Mittel vergangenheitsbezogen ist. Hierdurch zeigt sich wiederum, dass es aufgrund der qualitativen Faktoren wie sie bei der FondsNote verwendet werden, zu einer Verbesserung der Vorhersagekraft der Fondsbewertung kommt.

5.3 Einfluss des Fondsvermögens

Tabelle 7 und Abbildung A1 stellen die Beziehung zwischen der Veränderung des Fondsvermögens und den drei Fondsbewertungen dar. Sowohl die Zuordnung der Fonds in Bewertungsklassen basierend auf dem Feri Trust Rating, der Finanztest-Bewertung als auch der FondsNote führen zu einem ähnlichen Investitionsverhalten.

Fonds mit Top-Bewertung (Q5) weisen für den Gesamtzeitraum und der Marktphase mit steigenden Kursen in der Regel überproportionale Mittelzuflüsse auf. Nur bei Fonds mit FondsNote sind die Mittelzuflüsse geringer, was auf die geringere Bekanntheit dieser Fondsbewertung zurückgeführt werden kann. In Marktphasen mit fallenden Kursen dreht sich diese Beziehung komplett um. Fonds mit Top-Bewertung verzeichnen nun die höchsten Mittelabflüsse. Demgegenüber ist bei niedriger bewerteten Fonds (Q1 bis Q4) nur eine schwache Beziehung zwischen der Veränderung des Fondsvermögens und der Fondsbewertung festzustellen. Diese Fonds verzeichnen in der Regel geringe Mittelabflüsse, unabhängig davon, ob es sich um eine Marktphase mit steigenden oder fallenden Kursen handelt.

Interessanterweise unterscheiden sich die Mittelflüsse zwischen den am niedrigsten und am höchsten bewerteten Fonds nur beim Feri Trust Rating und der Finanztest-Bewertung hochsignifikant voneinander, wenn der Gesamtzeitraum oder der Zeitraum mit steigenden Kursen zugrunde liegt. Bei sinkenden Kursen ist jedoch bei allen drei Fondsbewertungen kein signifikanter Unterschied zu beobachten. Dies deutet darauf hin, dass Investoren vor allem bei steigenden Kursen Fondsbewertungen wie das Feri Trust Rating und die Finanztest-Bewertung in ihren Entscheidungsprozess mit einbeziehen und entsprechend die am höchsten bewerteten Fonds kaufen. In Marktphasen, in denen die Kurse sinken, werden stattdessen Fonds unabhängig von ihrer Fondsbewertung verkauft.

Tabelle 7: Fondsbewertung und Veränderung des Fondsvermögens

	Q1	Q2	Q3	Q4	Q5	Q5-Q1	T-Stat.
Feri Trust							
Gesamtzeitraum	-1.38	-1.44	-3.33	-3.12	7.25	**8.64*** **	(3.298)
12/2002-06/2007	-1.44	-1.35	-4.60	-3.98	9.68	**11.12*** **	(3.492)
12/2007-02/2009	-1.61	-1.83	-2.99	-4.24	-3.45	-1.84	(-0.303)
Finanztest							
Gesamtzeitraum	-1.96	-1.57	-3.98	-2.14	6.11	**8.07*** **	(3.934)
12/2002-06/2007	-2.04	-1.29	-5.54	-2.49	9.16	**11.20*** **	(5.868)
12/2007-02/2009	-1.28	-1.66	-3.40	-3.51	-8.72	-7.44	(-0.801)
FondsNote							
Gesamtzeitraum	-1.15	-2.41	-2.85	0.68	1.19	2.33	(1.274)
12/2002-06/2007	-1.68	-3.13	-3.99	2.34	2.62	**4.31***	(1.820)
12/2007-02/2009	-0.17	-1.56	-1.47	-1.54	-6.19	-6.02	(-1.256)

Die Tabelle stellt die Beziehung zwischen den Bewertungsklassen (Q1 bis Q5) und der Veränderung des Fondsvermögens dar. Beispielsweise entspricht Q1 dem arithmetischen Mittel der monatlichen Veränderung des Fondsvermögens der am niedrigsten bewerteten Fonds. Q5 entspricht dem arithmetischen Mittel der monatlichen Veränderung des Fondsvermögens von Fonds mit Top-Bewertung. Die Veränderung des Fondsvermögens ist in Mio. EUR angegeben und für jeweils drei Zeiträume berechnet worden (Gesamtzeitraum: 12/2002-12/2009; Bullenmarkt: 12/2002-06/2007; Bärenmarkt: 12/2007-02/2009). *, ** und *** zeigen die Signifikanz des zweiseitigen T-Tests auf dem 10%, 5% und 1% Niveau an.

Die hohen Mittelzuflüsse in Marktphasen mit steigenden Kursen können dazu führen, dass Top-Fonds mit Feri Trust Rating und Finanztest-Bewertung ihren Performancevorteil gegenüber geringer bewerteten Fonds verlieren. Dies kann ein Grund dafür sein, dass die Vorhersagekraft gerade bei diesen beiden Fondsbewertungen geringer ist als bei der FondsNote. Des Weiteren können die hohen Mittelzuflüsse bei Fonds mit Top-Bewertung in Marktphasen mit steigenden Kursen und die gleichermaßen hohen Mittelabflüsse bei sinkenden Kursen auf Momentum-Strategien und die Realisierung von Erträgen zurückgeführt werden. Die nur geringen Veränderungen im Fondsvermögen bei niedriger bewerteten Fonds, unabhängig von der Marktphase, zeigen, dass Investoren mögliche Performanceverluste niedrig bewerteter Fonds in der Hoffnung auf zukünftige Erträge kaum realisieren.

5.4 Einfluss des Total Expense Ratios

Als letztes soll der Frage nachgegangen werden, wie stark die Fondskosten die Vorhersagekraft einer Fondsbewertung beeinflussen. Tabelle 8 und Abbildung A2 stellen die Ergebnisse des Feri Trust Ratings, der Finanztest-Bewertung und der FondsNote einander gegenüber. Zudem werden die neu gebildeten Bewertungsklassen, die auf dem 4-Faktor-Alpha und dem geometrischen Mittel basieren, analysiert.

Die Ergebnisse zeigen, dass Fonds mit der dritthöchsten Fondsbewertung die geringsten Kosten besitzen. Sowohl Fonds in höheren als auch niedrigeren Bewertungsklassen weisen höhere

Kosten auf, wobei jene Fonds mit der höchsten und der niedrigsten Bewertung die höchsten Kosten haben. Dies kann teilweise die in der Literatur gefundene Erkenntnis erklären, weshalb sich die Performance der am höchsten bewerteten Fonds kaum von der der Zweithöchsten unterscheidet. Allerdings ist der Kostenunterschied zwischen den Bewertungsklassen (Q1 bis Q5) in der Regel geringer als der dazugehörige Performanceunterschied. Die geringe Vorhersagekraft des Feri Trust Ratings und der Finanztest-Bewertung und die etwas höhere Vorhersagekraft der FondsNote kann daher nur teilweise mithilfe der Fondskosten erklärt werden.

Tabelle 8: Fondsbewertung und Fondskosten

	Q1	Q2	Q3	Q4	Q5	Q5-Q1	T-Stat.
Feri Trust	0.0149	0.0136	0.0128	0.0132	0.0145	-0.0004	(-0.484)
Finanztest	0.0146	0.0138	0.0139	0.0145	0.0144	-0.0002	(-0.213)
FondsNote	0.0163	0.0142	0.0136	0.0140	0.0150	-0.0013	(-1.207)
Feri Trust							
4-Faktor-Alpha	0.0149	0.0132	0.0125	0.0135	0.0147	-0.0002	(-0.192)
GeoMittel	0.0150	0.0130	0.0127	0.0138	0.0140	-0.0010	(-1.330)
Finanztest							
4-Faktor-Alpha	0.0160	0.0138	0.0131	0.0149	0.0172	0.0012	(1.065)
GeoMittel	0.0160	0.0130	0.0136	0.0152	0.0172	0.0011	(1.049)
FondsNote							
4-Faktor-Alpha	0.0163	0.0136	0.0132	0.0148	0.0166	0.0002	(0.211)
GeoMittel	0.0155	0.0133	0.0133	0.0155	0.0159	0.0004	(0.302)

Die Tabelle stellt die Beziehung zwischen den Bewertungsklassen (Q1 bis Q5) der Fondsbewertungen (Feri Trust Rating, Finanztest-Bewertung, FondsNote, 4-Faktor-Alpha und geometrisches Mittel) und den Fondskosten dar. Beispielsweise entspricht Q1 dem arithmetischen Mittel der Fondskosten (annualisiertes Total Expense Ratio) der am niedrigsten bewerteten Fonds im Dezember eines jeden Jahres von 2005 bis 2009. Q5 stellt das arithmetische Mittel der Fondskosten von Fonds mit Top-Bewertung dar. *, ** und *** zeigen die Signifikanz des zweiseitigen T-Tests auf dem 10%, 5% und 1% Niveau an.

Zudem ist festzuhalten, dass sich die Kosten der Fonds mit Top-Bewertung (Q5) kaum von Fonds mit der niedrigsten Bewertung (Q1) unterscheiden. Ein Grund hierfür kann sein, dass Fonds mit der höchsten Bewertung gezielt vermarktet werden. Die hierbei entstehenden zusätzlichen Kosten für den Kauf der Vermarktungsrechte der Fondsbewertung und der damit verbundenen Werbung werden an die Investoren weitergegeben. Andererseits kann es auch sein, dass die mit einer Top-Bewertung verbundenen Mittelzuflüsse von der Fondsgesellschaft antizipiert werden. Um zusätzliche Erträge zu generieren, werden die Fondskosten erhöht. Dies setzt allerdings voraus, dass das Investitionsverhalten stärker von der Fondsbewertung als von den Fondskosten abhängt.

6 Zusammenfassung

In dieser Arbeit wird der Frage nachgegangen, ob es mithilfe von Fondsbewertungen gelingt, die zukünftige Performance von Fonds vorherzusagen. Neben dem Feri Trust Rating und der Finanztest-Bewertung, deren Methodik auf quantitativen Faktoren basiert, wird die FondsNote, die sowohl quantitative als auch qualitative Faktoren mit einbezieht, analysiert. Dies ist insofern eine Neuerung, da in der Literatur bis dato fast ausnahmslos Fondsbewertungen mit quantitativen Faktoren untersucht wurden. Zudem ist dies meines Erachtens die erste Arbeit, die die Vorhersagekraft verschiedener Fondsbewertungen für einen identischen Zeitraum und Anlageschwerpunkt gegenüberstellt. Der Vergleich basiert auf einem Zeitraum von 2002 bis 2010 und bezieht sämtliche Aktienfonds mit Anlageschwerpunkt Deutschland mit ein. Als Performancemaße werden das 4-Faktor-Alpha und das geometrische Mittel verwendet.

Die Ergebnisse zeigen, dass das Feri Trust Rating und die Finanztest-Bewertung die Performance von Fonds kaum vorhersagen können. Dies gilt für kurzfristige Ein-, Drei- und auch langfristige Fünf-Jahres-Zeiträume. Der FondsNote gelingt die Differenzierung zwischen sich zukünftig besser und schlechter entwickelnden Fonds deutlich besser. Fonds mit der niedrigsten Bewertung erzielen häufiger eine signifikant geringere Performance als Fonds mit Top-Bewertung. Auch der Performanceunterschied zwischen den Bewertungsklassen ist häufiger signifikant positiv als mit dem Feri Trust Rating und der Finanztest-Bewertung.

Zudem wird untersucht, ob eine Fondsbewertung eine höhere Vorhersagekraft besitzt, wenn sie auf dem 4-Faktor-Alpha oder dem geometrischen Mittel basiert. Interessanterweise führt der Vergleich der alternativen Fondsbewertungen im Vergleich zur FondsNote zu einer Verringerung der Differenzierung zwischen sich zukünftig besser und schlechter entwickelnden Fonds. Hierdurch kann gezeigt werden, dass die höhere Vorhersagekraft bei der FondsNote auf die Bewertung der Fonds mit qualitativen Faktoren zurückzuführen ist. Im Vergleich zum Feri Trust Rating ist mit den beiden Performancemaßen keine Veränderung in der Vorhersagekraft zu verzeichnen. Demgegenüber erhöht sich die Vorhersagekraft der Performancemaße im Vergleich zur Finanztest-Bewertung sogar.

Letztlich wird in dieser Arbeit die Beziehung zwischen Fondsvermögen und Fondsbewertung und zwischen Fondskosten und Fondsbewertung untersucht. Die Ergebnisse zeigen, dass die drei Fondsbewertungen das Fondsvermögen und somit das Investitionsverhalten beeinflussen. Insbesondere im Bullenmarkt verzeichnen Fonds mit Top-Bewertung überproportionale Mittelzuflüsse. Starke Mittelabflüsse sind hingegen im Bärenmarkt zu beobachten. Dies gilt vor allem für Fonds mit Feri Trust Rating und Finanztest-Bewertung, wodurch sich die geringe Vorhersagekraft dieser

Fondsbewertungen teilweise erklären lässt. Der Vergleich von Fondskosten und Fondsbewertung zeigt, dass insbesondere die am höchsten und am niedrigsten bewerteten Fonds die höchsten Kosten aufweisen. Die Fondskosten können allerdings nur bedingt die geringe Vorhersagekraft der Fondsbewertungen erklären.

Literaturverzeichnis

Antypas, A., Caporale, G. M., Kourogenis, N., Pittas, N. (2009). Selectivity, market timing and the Morningstar star-rating system. Diskussionspapier, Nr. 874, DIW Berlin, April.

Axel Springer, FondsConsult (2007). FondsNote. Das Fonds-Rating von Axel Springer Finanzen Verlag und FondsConsult. Axel Springer Finanzen Verlag & FondsConsult Research AG, München.

Blake, C. R., Morey, M. R. (2000). Morningstar ratings and mutual fund performance. The Journal of Financial and Quantitative Analysis 35 (3), S. 451-483.

Brückner, R., Lehmann, P., Schmidt, M. H., Stehle, R. (2014). Fama/French factors for Germany: which set is best? Diskussionspapier, Humboldt-Universität zu Berlin, Februar.

Del Guercio, D., Tkac, P. A. (2008). Star power: the effect of Morningstar ratings on mutual fund flow. The Journal of Financial and Quantitative Analysis 43 (4), S. 907-936.

Duret, A., Hereil, P., Mitaine, P., Moussavi, N., Roncalli, T. (2008). Fund rating systems and performance predictability. Diskussionspapier, Universität Évry, April.

Feri Trust (2002). Feri Trust Fonds Rating im Detail. Feri Trust GmbH, Juni.

Füss, R., Hille, J., Rindler, P., Schmidt, J., Schmidt, M. (2010). From rising stars and falling angels: on the relationship between the performance and ratings of German mutual funds. Journal of Wealth Management 13 (1), S. 75-90.

Garnier, O., Pujol, T. (2007). Can today's stars be used to read the stars of the future? Diskussionspapier, Nr. 3, März.

Gerrans, P. (2006). Morningstar ratings and future performance. Accounting and Finance 46 (4), S. 605-628.

Hereil, P., Mitaine, P., Moussavi, N., Roncalli, T. (2010). Mutual fund ratings and performance persistence. White Paper 3, Quant Research by Lyxor, Juni.

Huber, C. (2012). The effects of ratings on financial decision-making. In: Behavioral insurance: essays on the influence of ratings and price presentation on consumer evaluation, risk perception, and financial decision-making. Dissertation, Universität St. Gallen.

Jain, P. C., Wu, J. S. (2000). Truth in mutual fund advertising: evidence on future performance and fund flows. The Journal of Finance 55 (2), S. 937-958.

Kräussl, R., Sandelowsky, R. (2007). The predictive performance of Morningstar's mutual fund ratings. Diskussionspapier, Freie Universität Amsterdam, August.

McDonald, R. L, Rietz, T. A. (2014). Ratings and asset allocation: an experimental analysis. Diskussionspapier, Universität in Evanston & Universität in Iowa, Februar.

Meinhardt, C. (2009). Die Aussagekraft von Ratings und Rankings. Die Bank, Nr. 12, S. 14-17.

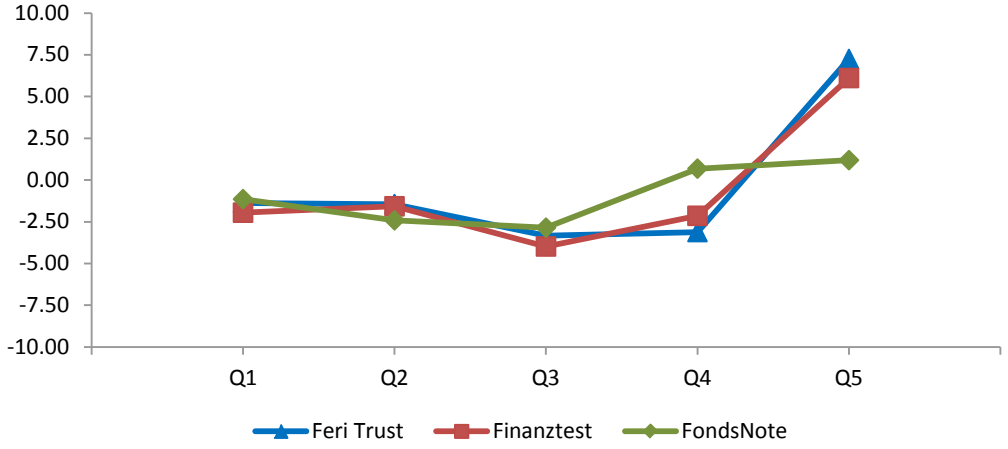
Morey, M. R., Gottesman, A. (2006). Morningstar mutual fund ratings redux. Journal of Investment Consulting 8 (1), S. 25-37.

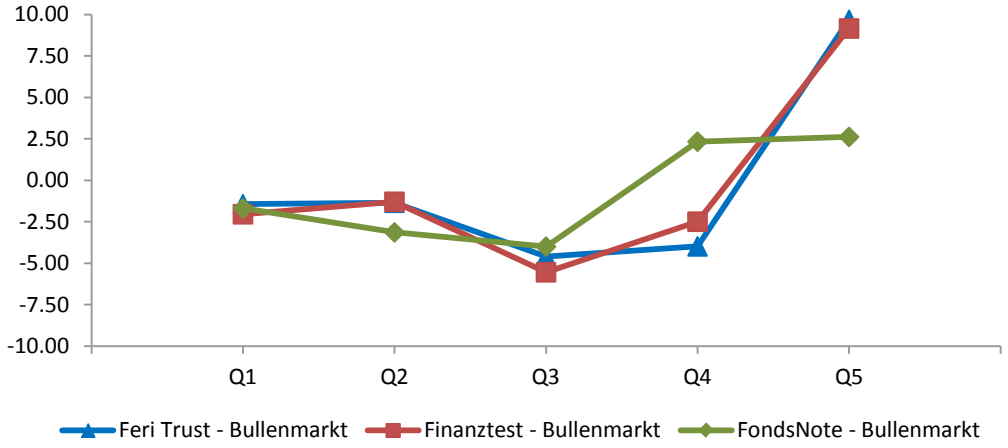
Müller, S., Weber, M. (2014). Evaluating the rating of Stiftung Warentest: how good are mutual fund ratings and can they be improved? European Financial Management 20 (2), S. 207-235.

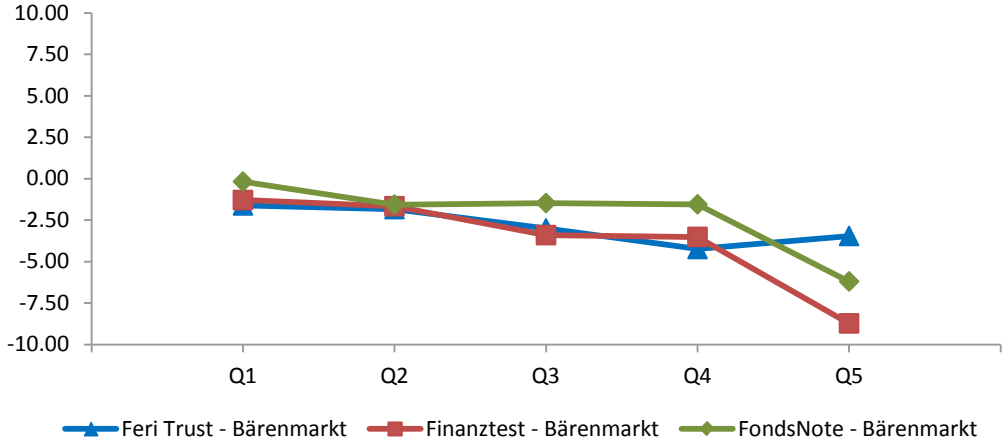
Sirri, E. R., Tufano, P. (1998). Costly search and mutual fund flows. The Journal of Finance 53 (5), S. 1589-1622.

Anhang

Abbildung A1: Fondsbewertung und Veränderung des Fondsvermögens

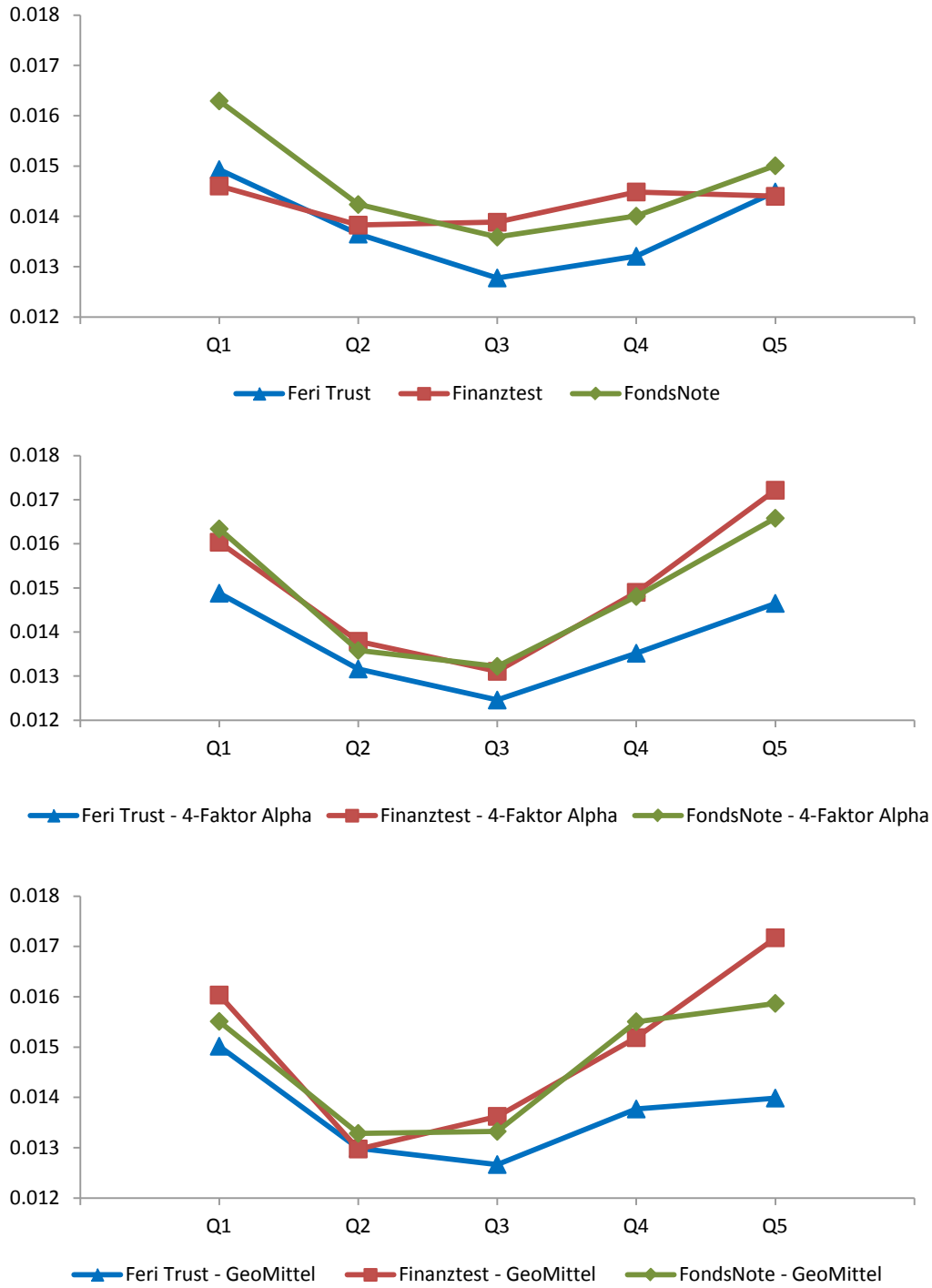






Die Abbildung stellt die Beziehung zwischen den Bewertungsklassen (Q1 bis Q5) der Fondsbewertungen (Feri Trust Rating, Finanztest-Bewertung und FondsNote) und der Veränderung des Fondsvermögens dar. Beispielsweise entspricht Q1 dem arithmetischen Mittel der monatlichen Veränderung des Fondsvermögens der am niedrigsten bewerteten Fonds. Q5 stellt das arithmetische Mittel der monatlichen Veränderung des Fondsvermögens von Fonds mit Top-Bewertung dar. Die Veränderung des Fondsvermögens ist in Mio. EUR angegeben und für jeweils drei Zeiträume berechnet worden (Gesamtzeitraum (oben): 12/2002-12/2009; Bullenmarkt (Mitte): 12/2002-06/2007; Bärenmarkt (unten): 12/2007-02/2009).

Abbildung A2: Fondsbewertung und Fondskosten



Die Abbildung stellt die Beziehung zwischen den Bewertungsklassen (Q1 bis Q5) der Fondsbewertungen und den Fondskosten dar. Beispielsweise entspricht Q1 dem arithmetischen Mittel der Fondskosten (annualisiertes Total Expense Ratio) der am niedrigsten bewerteten Fonds im Dezember eines jeden Jahres von 2005 bis 2009. Q5 stellt das arithmetische Mittel der Fondskosten von Fonds mit Top-Bewertung dar. Die erste Darstellung (oben) stellt das Feri Trust Rating, die Finanztest-Bewertung und die FondsNote einander gegenüber. In der zweiten Darstellung (Mitte) werden die Fondsbewertungen, die auf dem 4-Faktor-Alpha basieren, miteinander verglichen. Die dritte Darstellung (unten) bildet die Fondsbewertungen ab, die auf dem geometrischen Mittel basieren.

Tabelle A1.1: Feri Trust Rating

		4-Faktor-Alpha							Geometrisches Mittel						
Zeitraum		β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
1-Jahres Zeitraum	01/03-12/03	-0.011	-0.002	0.007	-0.009	0.020	0.03	0.643	**0.365*** **	-0.015	0.001	-0.016	0.016	0.04	0.833
		(-0.585)	(-0.100)	(0.362)	(-0.398)	(0.841)			(18.583)	(-0.635)	(0.027)	(-0.637)	(0.639)		
	01/04-12/04	**-0.084*** **	0.009	0.019	0.012	-0.001	0.05	0.949	**0.049*** **	0.001	0.015	0.004	0.012	0.04	0.794
		(-6.120)	(0.585)	(1.253)	(0.778)	(-0.089)			(3.807)	(0.058)	(1.017)	(0.292)	(0.764)		
	01/05-12/05	-0.057	-0.037	-0.022	-0.015	-0.022	0.01	0.242	**0.249*** **	-0.037	0.007	0.008	0.032	0.05	0.827
		(-1.552)	(-0.875)	(-0.513)	(-0.341)	(-0.449)			(6.078)	(-0.784)	(0.141)	(0.170)	(0.594)		
	01/06-12/06	**-0.077*** **	0.014	0.001	0.001	-0.023	0.07	1.161	**0.207*** **	0.005	0.003	0.016	**0.033* **	0.08	1.427
		(-4.781)	(0.748)	(0.068)	(0.070)	(-1.044)			(15.345)	(0.325)	(0.182)	(0.998)	(1.779)		
	01/07-12/07	**-0.046** **	-0.006	0.006	0.013	0.014	0.04	0.718	**0.158*** **	0.010	0.017	0.014	0.004	0.01	0.218
		(-2.542)	(-0.279)	(0.316)	(0.664)	(0.650)			(6.767)	(0.384)	(0.686)	(0.548)	(0.141)		
	01/08-12/08	0.038	0.053	**0.059* **	0.048	0.044	0.05	0.831	**-0.336*** **	**-0.080*** **	**-0.093*** **	**-0.091*** **	**-0.098*** **	0.22	**4.558*** **
		(1.311)	(1.507)	(1.787)	(1.249)	(1.134)			(-16.543)	(-3.222)	(-3.990)	(-3.343)	(-3.616)		
	01/09-12/09	**0.144*** **	**-0.077** **	**-0.079** **	**-0.086*** **	**-0.101*** **	0.13	**2.255* **	**0.325*** **	-0.055	**-0.066* **	**-0.067* **	**-0.075* **	0.07	1.096
		(5.160)	(-2.438)	(-2.456)	(-2.739)	(-2.758)			(10.365)	(-1.574)	(-1.834)	(-1.905)	(-1.838)		
	01/10-12/10	**-0.055** **	0.023	0.023	0.021	0.043	0.03	0.374	**0.150*** **	**0.046* **	0.045	0.035	**0.059* **	0.08	0.978
		(-2.217)	(0.826)	(0.820)	(0.770)	(1.212)			(6.132)	(1.699)	(1.651)	(1.336)	(1.687)		
3-Jahres Zeitraum	01/03-12/05	**-0.057*** **	-0.017	0.001	0.001	0.002	0.04	0.790	**0.222*** **	-0.034	-0.002	-0.006	0.006	0.08	1.643
		(-3.943)	(-0.984)	(0.067)	(0.039)	(0.130)			(11.685)	(-1.525)	(-0.092)	(-0.242)	(0.253)		
	01/04-12/06	-0.086	-0.045	0.008	0.006	0.001	0.03	0.481	**0.170*** **	-0.041	0.002	0.004	0.015	0.04	0.781
		(-1.333)	(-0.630)	(0.107)	(0.080)	(0.010)			(3.571)	(-0.785)	(0.038)	(0.078)	(0.253)		
	01/05-12/07	-0.063	-0.135	-0.011	-0.009	-0.006	0.04	0.600	**0.212*** **	-0.061	0.004	-0.002	0.014	0.04	0.762
		(-0.537)	(-0.987)	(-0.075)	(-0.066)	(-0.041)			(3.978)	(-0.987)	(0.066)	(-0.033)	(0.192)		
	01/06-12/08	-0.002	0.011	0.023	**0.031* **	0.019	0.07	1.148	**-0.060*** **	0.000	0.010	0.004	-0.012	0.05	0.809
		(-0.114)	(0.642)	(1.330)	(1.766)	(0.923)			(-5.129)	(-0.036)	(0.728)	(0.266)	(-0.734)		
	01/07-12/09	0.028	-0.012	0.020	0.033	0.016	0.06	0.994	-0.042	-0.019	-0.005	-0.006	-0.015	0.02	0.255
		(0.831)	(-0.322)	(0.543)	(0.864)	(0.390)			(-1.644)	(-0.646)	(-0.165)	(-0.196)	(-0.466)		
	01/08-12/10	-0.032	0.064	0.047	0.057	**0.079* **	0.06	0.911	**-0.075*** **	0.035	0.014	0.023	0.043	0.04	0.562
		(-0.966)	(1.541)	(1.224)	(1.303)	(1.799)			(-2.793)	(1.069)	(0.466)	(0.666)	(1.224)		
5-Jahres Zeitraum	01/03-12/07	-0.056	-0.081	-0.002	-0.004	0.001	0.04	0.699	**0.212*** **	-0.063	-0.003	-0.007	0.005	0.05	0.984
		(-0.811)	(-1.011)	(-0.028)	(-0.043)	(0.014)			(4.732)	(-1.210)	(-0.068)	(-0.124)	(0.083)		
	01/04-12/08	-0.043	-0.067	0.030	0.027	0.026	0.04	0.722	0.000	-0.024	0.028	0.021	0.016	0.04	0.728
		(-0.441)	(-0.617)	(0.269)	(0.246)	(0.220)			(0.005)	(-0.439)	(0.516)	(0.377)	(0.263)		
	01/05-12/09	0.008	-0.111	0.011	0.002	-0.004	0.05	0.874	0.053	-0.055	0.012	0.001	0.004	0.05	0.939
		(0.089)	(-1.086)	(0.099)	(0.019)	(-0.031)			(1.177)	(-1.052)	(0.229)	(0.027)	(0.071)		
	01/06-12/10	-0.007	-0.004	0.013	0.007	0.010	0.02	0.241	**0.040* **	-0.005	0.008	-0.003	0.007	0.01	0.187
		(-0.279)	(-0.128)	(0.476)	(0.237)	(0.297)			(1.797)	(-0.196)	(0.301)	(-0.112)	(0.215)		

Dummy-Variablen Regression mit 4-Faktor-Alpha und geometrischem Mittel. *, ** und *** zeigen die Signifikanz auf dem 10%, 5% und 1% Niveau an.

Tabelle A1.2: Finanztest-Bewertung

		4-Faktor-Alpha							Geometrisches Mittel						
Zeitraum		β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
1-Jahres	01/03-12/03	-0.042	0.046	0.062	0.081		0.04	1.295	**0.343*** **	0.034	0.038	0.055		0.02	0.458
Zeitraum		(-0.943)	(0.996)	(1.344)	(1.662)				(7.324)	(0.689)	(0.774)	(1.067)			
	01/04-12/04	**-0.080*** **	0.007	0.016	0.008	-0.019	0.06	1.136	**0.056*** **	0.002	0.013	0.024	0.002	0.04	0.769
		(-6.484)	(0.510)	(1.122)	(0.559)	(-0.913)			(3.311)	(0.079)	(0.653)	(1.187)	(0.079)		
	01/05-12/05	**-0.086*** **	0.017	0.013	0.032	-0.017	0.06	1.221	**0.258*** **	0.006	0.017	0.026	0.038	0.05	0.864
		(-5.653)	(0.946)	(0.647)	(1.527)	(-0.672)			(16.668)	(0.333)	(0.854)	(1.213)	(1.482)		
	01/06-12/06	**-0.102*** **	0.028	0.022	0.029	-0.020	0.08	1.735	**0.200*** **	0.006	0.024	**0.034** **	**0.062*** **	0.17	**3.736*** **
		(-6.389)	(1.536)	(1.245)	(1.439)	(-0.783)			(15.845)	(0.434)	(1.654)	(2.124)	(3.022)		
	01/07-12/07	0.016	**-0.069*** **	**-0.052*** **	**-0.053*** **	-0.024	0.17	**3.946*** **	**0.155*** **	0.004	0.008	-0.016	0.013	0.02	0.475
		(0.918)	(-3.630)	(-2.801)	(-2.676)	(-0.932)			(5.769)	(0.145)	(0.260)	(-0.512)	(0.332)		
	01/08-12/08	**0.107*** **	-0.039	-0.034	-0.008	-0.054	0.03	0.646	**-0.355*** **	**-0.059* **	**-0.069** **	**-0.067** **	-0.110	0.08	1.604
		(3.014)	(-0.983)	(-0.879)	(-0.213)	(-0.622)			(-13.075)	(-1.953)	(-2.346)	(-2.237)	(-1.647)		
	01/09-12/09	**0.121*** **	0.005	-0.051	-0.055	-0.047	0.10	**2.076* **	**0.303*** **	0.027	-0.034	-0.029	-0.034	0.06	1.343
		(3.652)	(0.121)	(-1.430)	(-1.556)	(-0.879)			(7.882)	(0.592)	(-0.815)	(-0.715)	(-0.546)		
	01/10-12/10	0.011	-0.050	-0.033	0.008	-0.061	0.06	1.123	**0.197*** **	-0.048	0.006	0.060	-0.027	0.14	**2.711** **
		(0.228)	(-0.924)	(-0.673)	(0.162)	(-0.736)			(4.126)	(-0.876)	(0.131)	(1.136)	(-0.328)		
3-Jahres	01/03-12/05	**-0.069*** **	0.015	0.021	0.027		0.03	0.835	**0.211*** **	0.016	0.023	0.028		0.01	0.421
Zeitraum		(-3.704)	(0.800)	(1.076)	(1.336)				(7.678)	(0.574)	(0.801)	(0.932)			
	01/04-12/06	**-0.095*** **	0.014	0.017	0.017	-0.011	0.09	1.761	**0.168*** **	0.002	0.020	0.026	0.017	0.08	1.542
		(-10.236)	(1.320)	(1.598)	(1.560)	(-0.742)			(12.223)	(0.155)	(1.321)	(1.566)	(0.764)		
	01/05-12/07	**-0.072*** **	-0.004	0.001	0.001	-0.009	0.02	0.301	**0.208*** **	0.000	0.013	0.003	0.014	0.05	0.866
		(-10.103)	(-0.494)	(0.095)	(0.054)	(-0.748)			(24.942)	(-0.037)	(1.217)	(0.219)	(1.030)		
	01/06-12/08	-0.004	0.009	**0.027* **	**0.030* **	**0.043** **	0.11	**2.275* **	**-0.084*** **	0.021	**0.026* **	0.021	0.007	0.05	1.089
		(-0.286)	(0.654)	(1.916)	(1.928)	(2.162)			(-6.890)	(1.474)	(1.899)	(1.378)	(0.356)		
	01/07-12/09	**0.079*** **	**-0.040** **	**-0.038* **	-0.011	-0.013	0.12	**2.543** **	**-0.037*** **	-0.013	-0.010	-0.011	-0.033	0.03	0.679
		(4.400)	(-2.058)	(-1.946)	(-0.536)	(-0.466)			(-2.704)	(-0.836)	(-0.678)	(-0.685)	(-1.588)		
	01/08-12/10	**0.071*** **	**-0.044* **	-0.033	-0.031	0.005	0.06	1.070	-0.015	-0.022	-0.020	-0.022	-0.015	0.01	0.265
		(3.238)	(-1.828)	(-1.433)	(-1.297)	(0.099)			(-0.715)	(-0.986)	(-0.905)	(-0.970)	(-0.314)		
5-Jahres	01/03-12/07	**-0.064*** **	0.007	0.010	0.014		0.02	0.575	**0.202*** **	0.007	0.013	0.017		0.02	0.686
Zeitraum		(-5.162)	(0.512)	(0.794)	(0.999)				(12.940)	(0.424)	(0.786)	(0.974)			
	01/04-12/08	**-0.044*** **	**0.023** **	**0.025** **	**0.030*** **	0.023	0.10	**2.139* **	-0.014	**0.032*** **	**0.036*** **	**0.039*** **	0.022	0.15	**3.430*** **
		(-5.007)	(2.321)	(2.565)	(2.853)	(1.561)			(-1.459)	(2.938)	(3.302)	(3.461)	(1.373)		
	01/05-12/09	0.011	-0.004	-0.001	-0.009	0.008	0.03	0.488	**0.053*** **	0.004	0.006	0.010	0.006	0.02	0.333
		(1.430)	(-0.456)	(-0.123)	(-0.799)	(0.609)			(8.008)	(0.476)	(0.697)	(1.101)	(0.558)		
	01/06-12/10	-0.003	0.006	0.011	0.016	0.025	0.04	0.672	**0.030*** **	0.016	**0.017* **	0.019	**0.027* **	0.05	1.035
		(-0.265)	(0.477)	(0.881)	(1.114)	(1.387)			(3.268)	(1.505)	(1.683)	(1.643)	(1.793)		

Dummy-Variablen Regression mit 4-Faktor-Alpha und geometrischem Mittel. *, ** und *** zeigen die Signifikanz auf dem 10%, 5% und 1% Niveau an.

Tabelle A1.3: FondsNote

		4-Faktor-Alpha							Geometrisches Mittel						
Zeitraum		β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
1-Jahres Zeitraum	01/03-12/03	**-0.105***	**0.107***	**0.113****	**0.120****	**0.167*****	0.15	**2.398****	**0.354*****	0.012	0.018	0.028	0.063	0.04	0.620
		(-1.940)	(1.911)	(2.043)	(2.137)	(2.816)			(5.237)	(0.168)	(0.257)	(0.395)	(0.851)		
	01/04-12/04	**-0.118*****	**0.050****	**0.047****	**0.053****	**0.079*****	0.14	**2.450****	0.027	**0.050****	0.027	**0.047***	**0.109*****	0.23	**4.405*****
		(-6.394)	(2.401)	(2.360)	(2.558)	(3.003)			(1.283)	(2.073)	(1.181)	(1.983)	(3.614)		
	01/05-12/05	**-0.133*****	**0.077****	0.054	0.056	0.039	0.08	1.431	**0.289*****	-0.035	-0.012	-0.011	-0.007	0.04	0.628
		(-4.150)	(2.192)	(1.621)	(1.578)	(0.953)			(9.302)	(-1.015)	(-0.376)	(-0.316)	(-0.167)		
	01/06-12/06	**-0.111*****	**0.035****	**0.039****	0.015	-0.007	0.16	**3.321*****	**0.219*****	-0.012	-0.003	0.025	0.019	0.16	**3.444*****
		(-6.879)	(1.996)	(2.146)	(0.837)	(-0.300)			(14.006)	(-0.720)	(-0.164)	(1.394)	(0.914)		
	01/07-12/07	-0.019	0.003	-0.025	-0.020	-0.002	0.08	1.429	**0.134*****	0.042	0.039	0.016	-0.019	0.14	**2.860****
		(-1.075)	(0.147)	(-1.274)	(-1.000)	(-0.103)			(5.854)	(1.581)	(1.537)	(0.612)	(-0.669)		
	01/08-12/08	0.005	**0.099****	**0.082***	**0.077***	**0.085***	0.06	1.255	**-0.448*****	0.043	0.012	0.022	0.011	0.06	1.145
		(0.126)	(2.200)	(1.848)	(1.715)	(1.733)			(-14.229)	(1.269)	(0.353)	(0.660)	(0.299)		
	01/09-12/09	**0.280*****	**-0.134*****	**-0.158*****	**-0.214*****	**-0.150*****	0.26	**7.358*****	**0.483*****	**-0.176*****	**-0.164*****	**-0.219*****	**-0.144****	0.20	**5.464*****
		(6.044)	(-2.685)	(-3.216)	(-4.438)	(-2.645)			(9.551)	(-3.224)	(-3.048)	(-4.158)	(-2.315)		
	01/10-12/10	0.015	-0.030	-0.057	-0.025	-0.033	0.04	0.873	**0.255*****	-0.047	-0.066	-0.052	**-0.081***	0.04	0.924
		(0.376)	(-0.694)	(-1.358)	(-0.609)	(-0.679)			(6.681)	(-1.100)	(-1.621)	(-1.288)	(-1.696)		
3-Jahres Zeitraum	01/03-12/05	**-0.145*****	**0.089*****	**0.089*****	**0.097*****	**0.117*****	0.26	**4.742*****	**0.225*****	-0.009	-0.002	0.009	0.020	0.09	1.298
		(-5.691)	(3.355)	(3.416)	(3.671)	(4.190)			(7.769)	(-0.299)	(-0.073)	(0.297)	(0.624)		
	01/04-12/06	**-0.151*****	**0.062*****	**0.066*****	**0.080*****	**0.082*****	0.27	**5.476*****	**0.162*****	0.029	0.014	0.026	**0.055****	0.10	1.604
		(-9.535)	(3.433)	(3.892)	(4.542)	(3.645)			(8.639)	(1.346)	(0.675)	(1.239)	(2.079)		
	01/05-12/07	**-0.084*****	0.014	0.012	0.020	0.004	0.06	1.023	**0.215*****	-0.006	0.008	-0.004	-0.022	0.11	**2.059***
		(-7.329)	(1.092)	(1.016)	(1.588)	(0.303)			(14.337)	(-0.370)	(0.507)	(-0.264)	(-1.176)		
	01/06-12/08	-0.005	0.025	0.023	**0.040****	0.029	0.08	1.458	**-0.129*****	**0.067*****	**0.069*****	**0.064*****	**0.037****	0.31	**8.054*****
		(-0.339)	(1.471)	(1.335)	(2.267)	(1.386)			(-10.509)	(4.978)	(5.054)	(4.621)	(2.247)		
	01/07-12/09	0.024	**0.052*****	0.024	0.027	**0.051****	0.15	**3.097****	**-0.054*****	-0.001	0.007	0.008	0.000	0.02	0.408
		(1.513)	(2.789)	(1.340)	(1.496)	(2.627)			(-4.971)	(-0.087)	(0.566)	(0.643)	(0.035)		
	01/08-12/10	0.039	-0.009	-0.005	0.005	0.029	0.07	1.398	**-0.096*****	**0.052***	0.043	**0.064****	**0.080*****	0.13	**2.614****
		(1.350)	(-0.309)	(-0.152)	(0.161)	(0.893)			(-3.808)	(1.916)	(1.617)	(2.380)	(2.732)		
5-Jahres Zeitraum	01/03-12/07	**-0.118*****	**0.059*****	**0.062*****	**0.069*****	**0.075*****	0.28	**5.132*****	**0.146*****	**0.061*****	**0.064*****	**0.076*****	**0.056****	0.21	**3.601*****
		(-7.193)	(3.516)	(3.714)	(4.092)	(4.206)			(6.948)	(2.787)	(2.968)	(3.461)	(2.446)		
	01/04-12/08	**-0.060*****	**0.045*****	**0.044*****	**0.043*****	**0.056*****	0.25	**5.030*****	**-0.035*****	**0.060*****	**0.054*****	**0.058*****	**0.074*****	0.36	**8.326*****
		(-5.962)	(3.955)	(4.050)	(3.837)	(3.978)			(-3.433)	(5.154)	(4.924)	(5.074)	(5.105)		
	01/05-12/09	0.008	-0.003	0.009	0.005	0.007	0.03	0.576	**0.020***	**0.035*****	**0.043*****	**0.036*****	**0.043*****	0.17	**3.398*****
		(0.509)	(-0.162)	(0.593)	(0.287)	(0.370)			(1.735)	(2.801)	(3.569)	(2.846)	(2.917)		
	01/06-12/10	0.002	0.006	0.003	0.022	**0.032***	0.11	**2.203***	-0.001	**0.045*****	**0.047*****	**0.055*****	**0.064*****	0.29	**7.273*****
		(0.128)	(0.408)	(0.213)	(1.387)	(1.776)			(-0.110)	(4.155)	(4.337)	(4.934)	(4.826)		

Dummy-Variablen Regression mit 4-Faktor-Alpha und geometrischem Mittel. *, ** und *** zeigen die Signifikanz auf dem 10%, 5% und 1% Niveau an.

Tabelle A2.1: Alternative Fondsbewertung – Dummy-Variablen Regression basierend auf Fonds mit Feri Trust Rating

		4-Faktor-Alpha							Geometrisches Mittel						
Zeitraum		β_1	β_2	β_3	β_4	β_5	R^2	**F-Stat.**	β_1	β_2	β_3	β_4	β_5	R^2	**F-Stat.**
1-Jahres Zeitraum	01/03-12/03	-0.008	0.001	-0.002	-0.002	0.013	0.01	0.189	**0.368*** **	-0.015	-0.011	0.001	0.010	0.03	0.534
		(-0.400)	(0.026)	(-0.090)	(-0.102)	(0.543)			(18.735)	(-0.645)	(-0.496)	(0.039)	(0.407)		
	01/04-12/04	**-0.074*** **	-0.003	0.005	0.008	-0.009	0.04	0.645	**0.049*** **	0.005	0.004	0.007	**0.025* **	0.06	1.073
		(-6.393)	(-0.229)	(0.390)	(0.586)	(-0.569)			(4.459)	(0.431)	(0.326)	(0.543)	(1.721)		
	01/05-12/05	**-0.081** **	-0.007	0.006	0.023	0.001	0.01	0.221	**0.249*** **	-0.040	0.009	0.022	0.007	0.05	0.840
		(-2.058)	(-0.155)	(0.129)	(0.500)	(0.012)			(5.564)	(-0.784)	(0.173)	(0.416)	(0.120)		
	01/06-12/06	**-0.075*** **	-0.003	0.008	0.012	-0.011	0.04	0.522	**0.211*** **	-0.001	0.011	0.009	0.016	0.04	0.524
		(-4.332)	(-0.142)	(0.376)	(0.572)	(-0.493)			(15.081)	(-0.063)	(0.675)	(0.518)	(0.899)		
	01/07-12/07	**-0.056*** **	**0.027* **	0.006	0.008	**0.040** **	0.14	**2.459** **	**0.162*** **	0.004	0.010	0.032	-0.022	0.12	**2.178* **
		(-4.005)	(1.672)	(0.398)	(0.462)	(2.203)			(8.949)	(0.206)	(0.474)	(1.488)	(-0.923)		
	01/08-12/08	0.017	**0.089** **	**0.079** **	**0.071* **	0.040	0.11	1.789	**-0.393*** **	-0.006	-0.032	-0.014	**-0.057* **	0.10	1.551
		(0.523)	(2.399)	(2.092)	(1.822)	(0.979)			(-16.598)	(-0.202)	(-1.156)	(-0.474)	(-1.855)		
	01/09-12/09	**0.063** **	-0.015	0.009	0.010	0.036	0.06	1.022	**0.292*** **	-0.007	-0.031	-0.056	-0.033	0.08	1.289
		(2.370)	(-0.502)	(0.278)	(0.307)	(1.057)			(10.286)	(-0.206)	(-0.928)	(-1.653)	(-0.894)		
	01/10-12/10	**-0.073*** **	**0.048** **	0.029	**0.039* **	**0.075*** **	0.21	**3.170** **	**0.161*** **	**0.038* **	0.014	**0.037* **	**0.048* **	0.13	1.735
		(-4.200)	(2.352)	(1.408)	(1.884)	(3.288)			(8.744)	(1.754)	(0.634)	(1.702)	(1.967)		
3-Jahres Zeitraum	01/03-12/05	**-0.055*** **	-0.011	-0.005	0.000	-0.004	0.01	0.222	**0.221*** **	-0.009	-0.027	0.002	0.008	0.06	1.079
		(-3.651)	(-0.650)	(-0.279)	(0.005)	(-0.207)			(11.167)	(-0.392)	(-1.183)	(0.104)	(0.298)		
	01/04-12/06	-0.084	-0.057	0.008	0.005	0.001	0.04	0.647	**0.172*** **	-0.004	-0.048	0.002	0.021	0.05	0.864
		(-1.510)	(-0.886)	(0.124)	(0.080)	(0.011)			(4.179)	(-0.093)	(-1.001)	(0.037)	(0.396)		
	01/05-12/07	-0.066	-0.135	0.000	-0.002	-0.012	0.04	0.624	**0.212*** **	-0.061	0.002	0.005	0.001	0.04	0.691
		(-0.515)	(-0.921)	(-0.001)	(-0.014)	(-0.073)			(3.631)	(-0.910)	(0.032)	(0.069)	(0.011)		
	01/06-12/08	-0.013	**0.037** **	**0.033* **	**0.038** **	0.026	0.09	1.344	**-0.043*** **	-0.016	-0.012	-0.014	-0.019	0.03	0.463
		(-0.882)	(2.126)	(1.892)	(2.071)	(1.347)			(-3.622)	(-1.187)	(-0.856)	(-0.971)	(-1.250)		
	01/07-12/09	0.016	0.036	0.021	0.039	0.019	0.03	0.483	**-0.056** **	0.004	0.010	0.007	-0.005	0.01	0.147
		(0.556)	(1.101)	(0.632)	(1.163)	(0.515)			(-2.631)	(0.177)	(0.402)	(0.260)	(-0.200)		
	01/08-12/10	-0.049	**0.090** **	**0.077* **	**0.077* **	0.057	0.08	1.282	**-0.095*** **	0.046	0.045	0.056	0.053	0.05	0.721
		(-1.379)	(2.172)	(1.841)	(1.781)	(1.242)			(-3.302)	(1.384)	(1.336)	(1.612)	(1.426)		
5-Jahres Zeitraum	01/03-12/07	-0.054	-0.081	-0.007	0.000	-0.007	0.04	0.612	**0.210*** **	-0.009	-0.062	0.002	0.006	0.05	0.875
		(-0.765)	(-0.980)	(-0.083)	(0.005)	(-0.079)			(4.560)	(-0.169)	(-1.143)	(0.028)	(0.105)		
	01/04-12/08	-0.021	-0.102	0.004	0.007	-0.001	0.05	0.809	0.025	-0.005	-0.055	-0.004	-0.005	0.04	0.735
		(-0.245)	(-1.043)	(0.036)	(0.069)	(-0.012)			(0.588)	(-0.096)	(-1.119)	(-0.089)	(-0.097)		
	01/05-12/09	0.014	-0.115	0.002	-0.002	-0.016	0.05	0.803	0.066	-0.070	-0.004	-0.009	-0.014	0.06	0.978
		(0.141)	(-1.044)	(0.021)	(-0.022)	(-0.131)			(1.337)	(-1.243)	(-0.071)	(-0.155)	(-0.220)		
	01/06-12/10	**-0.057** **	**0.073*** **	**0.063** **	**0.052* **	**0.059** **	0.13	**2.143* **	**0.057*** **	-0.023	-0.010	-0.023	-0.017	0.02	0.337
		(-2.604)	(2.856)	(2.453)	(1.949)	(2.074)			(2.744)	(-0.948)	(-0.419)	(-0.910)	(-0.614)		

Alternative Fondsbewertung mithilfe von Performancemaßen: 4-Faktor-Alpha und geometrisches Mittel (in-sample: 5 Jahre). *, ** und *** zeigen die Signifikanz auf dem 10%, 5% und 1% Niveau an.

Tabelle A2.2: Alternative Fondsbewertung – Dummy-Variablen Regression basierend auf Fonds mit Finanztest-Bewertung

		4-Faktor-Alpha							Geometrisches Mittel						
Zeitraum		β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
1-Jahres Zeitraum	01/03-12/03	**0.092*** **	**-0.058***	**-0.062** **	-0.048	-0.060	0.05	1.116	**0.471*** **	**-0.081** **	**-0.071** **	-0.015	-0.053	0.13	**2.963** **
		(3.394)	(-1.897)	(-2.029)	(-1.441)	(-1.379)			(16.832)	(-2.557)	(-2.247)	(-0.440)	(-1.164)		
	01/04-12/04	**-0.046*** **	-0.008	-0.001	-0.004	**-0.041** **	0.07	1.385	**0.081*** **	-0.001	0.000	0.022	**0.078*** **	0.16	**3.529*** **
		(-3.896)	(-0.606)	(-0.040)	(-0.301)	(-2.015)			(5.258)	(-0.072)	(-0.025)	(1.154)	(2.913)		
	01/05-12/05	**-0.102*** **	**0.063*** **	**0.056*** **	**0.065*** **	-0.038	0.27	**6.396*** **	**0.277*** **	0.015	0.016	0.040	**0.074** **	0.09	1.781
		(-5.762)	(3.142)	(2.797)	(2.952)	(-1.283)			(13.658)	(0.655)	(0.698)	(1.588)	(2.199)		
	01/06-12/06	**-0.085*** **	0.032	**0.042** **	0.033	0.017	0.06	1.174	**0.235*** **	0.011	0.019	**0.052*** **	0.038	0.16	**3.265** **
		(-4.727)	(1.543)	(2.070)	(1.472)	(0.571)			(16.333)	(0.680)	(1.198)	(2.903)	(1.584)		
	01/07-12/07	0.008	-0.003	-0.015	0.006	0.004	0.04	0.756	**0.199*** **	0.018	**0.036***	**-0.054** **	**-0.109*** **	0.43	**13.952*** **
		(0.566)	(-0.180)	(-0.909)	(0.339)	(0.168)			(11.680)	(0.895)	(1.846)	(-2.507)	(-3.703)		
	01/08-12/08	**0.112*** **	0.008	0.031	0.022	-0.023	0.03	0.486	**-0.473*** **	**0.114*** **	**0.072** **	**0.077** **	**0.111** **	0.23	**4.219*** **
		(3.456)	(0.209)	(0.837)	(0.542)	(-0.404)			(-18.534)	(3.927)	(2.471)	(2.437)	(2.507)		
	01/09-12/09	**0.175*** **	**-0.077** **	**-0.098*** **	**-0.114*** **	**-0.085** **	0.19	**3.591*** **	**0.371*** **	**-0.076** **	**-0.092*** **	**-0.090** **	-0.069	0.11	1.900
		(6.870)	(-2.643)	(-3.360)	(-3.546)	(-2.012)			(12.299)	(-2.190)	(-2.657)	(-2.364)	(-1.383)		
	01/10-12/10	**-0.045***	0.028	0.026	**0.066** **	**0.081***	0.10	1.769	**0.211*** **	-0.022	-0.001	0.021	**0.083***	0.12	**2.243***
		(-1.791)	(0.960)	(0.898)	(2.075)	(1.939)			(8.039)	(-0.723)	(-0.039)	(0.640)	(1.906)		
3-Jahres Zeitraum	01/03-12/05	-0.013	-0.014	**-0.022***	-0.011	-0.021	0.04	0.839	**0.286*** **	**-0.039** **	**-0.039** **	-0.003	-0.007	0.12	**2.844** **
		(-1.095)	(-1.097)	(-1.676)	(-0.785)	(-1.151)			(17.272)	(-2.080)	(-2.075)	(-0.128)	(-0.246)		
	01/04-12/06	**-0.067*** **	0.014	0.012	0.009	-0.014	0.07	1.428	**0.200*** **	-0.003	0.010	**0.031***	**0.044** **	0.14	**2.986** **
		(-7.600)	(1.373)	(1.138)	(0.821)	(-0.929)			(15.582)	(-0.213)	(0.678)	(1.897)	(1.997)		
	01/05-12/07	**-0.060*** **	**0.020***	**0.022** **	0.010	-0.008	0.14	**2.742** **	**0.235*** **	0.009	0.020	0.018	-0.007	0.08	1.444
		(-6.821)	(1.983)	(2.268)	(0.929)	(-0.536)			(21.290)	(0.719)	(1.623)	(1.343)	(-0.390)		
	01/06-12/08	0.014	**0.043*** **	**0.043*** **	**0.047*** **	**0.038***	0.13	**2.469** **	**-0.075*** **	**0.055*** **	**0.049*** **	**0.047*** **	0.028	0.19	**4.030*** **
		(1.079)	(2.831)	(2.859)	(2.850)	(1.720)			(-5.891)	(3.822)	(3.398)	(2.976)	(1.335)		
	01/07-12/09	**0.069*** **	0.019	0.007	0.015	0.004	0.03	0.490	**-0.027** **	0.008	0.015	-0.003	0.009	0.04	0.838
		(4.654)	(1.108)	(0.388)	(0.803)	(0.168)			(-2.401)	(0.636)	(1.136)	(-0.220)	(0.482)		
	01/08-12/10	**0.063*** **	-0.006	-0.006	-0.001	-0.031	0.03	0.354	**-0.055*** **	**0.049*** **	0.028	**0.039** **	**0.093*** **	0.22	**3.929*** **
		(3.769)	(-0.336)	(-0.298)	(-0.045)	(-1.070)			(-3.547)	(2.808)	(1.582)	(2.063)	(3.483)		
5-Jahres Zeitraum	01/03-12/07	**-0.022*** **	-0.004	-0.009	-0.003	-0.017	0.04	0.760	**0.252*** **	-0.014	-0.009	0.006	0.002	0.07	1.585
		(-2.915)	(-0.438)	(-0.986)	(-0.267)	(-1.408)			(25.769)	(-1.306)	(-0.804)	(0.472)	(0.102)		
	01/04-12/08	0.009	-0.009	0.005	0.010	-0.008	0.10	**2.008***	**0.026*** **	**0.019***	**0.025** **	**0.029** **	**0.045*** **	0.13	**2.574** **
		(1.083)	(-0.960)	(0.565)	(0.939)	(-0.552)			(2.789)	(1.786)	(2.317)	(2.520)	(2.788)		
	01/05-12/09	**0.045*** **	-0.015	0.000	-0.015	-0.011	0.07	1.351	**0.068*** **	**0.022** **	**0.021** **	**0.022** **	0.022	0.08	1.472
		(4.454)	(-1.307)	(-0.028)	(-1.198)	(-0.678)			(8.035)	(2.254)	(2.238)	(2.062)	(1.597)		
	01/06-12/10	0.017	0.021	0.015	0.021	0.013	0.04	0.727	**0.043*** **	**0.037*** **	**0.028** **	**0.036*** **	**0.033** **	0.16	**3.281** **
		(1.466)	(1.569)	(1.128)	(1.480)	(0.690)			(4.519)	(3.456)	(2.599)	(3.099)	(2.130)		

Alternative Fondsbewertung mithilfe von Performancemaßen: 4-Faktor-Alpha und geometrisches Mittel (in-sample: 5 Jahre). *, ** und *** zeigen die Signifikanz auf dem 10%, 5% und 1% Niveau an.

Tabelle A2.3: Alternative Fondsbewertung – Dummy-Variablen Regression basierend auf Fonds mit FondsNote

		4-Faktor-Alpha							Geometrisches Mittel						
Zeitraum		β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
1-Jahres	01/03-12/03	0.017	-0.011	0.000	-0.020	-0.005	0.03	0.373	**0.355*** **	0.012	0.017	0.021	**0.084***	0.09	1.301
Zeitraum		(0.607)	(-0.349)	(0.012)	(-0.628)	(-0.142)			(9.311)	(0.291)	(0.419)	(0.486)	(1.742)		
	01/04-12/04	**-0.060*** **	-0.004	-0.004	-0.002	**-0.041***	0.10	1.525	**0.046** **	0.014	0.016	0.028	**0.072** **	0.12	**2.055***
		(-3.179)	(-0.204)	(-0.201)	(-0.103)	(-1.711)			(2.045)	(0.538)	(0.649)	(1.090)	(2.362)		
	01/05-12/05	**-0.225*** **	**0.164*** **	**0.158*** **	**0.172*** **	**0.112*** **	0.45	**11.894*** **	**0.293*** **	-0.018	-0.026	-0.024	0.000	0.03	0.449
		(-9.155)	(6.127)	(6.038)	(6.388)	(3.618)			(10.841)	(-0.600)	(-0.891)	(-0.781)	(0.003)		
	01/06-12/06	**-0.144*** **	**0.057*** **	**0.060*** **	**0.077*** **	**0.056** **	0.18	**3.370*** **	**0.230*** **	-0.028	-0.012	0.008	0.016	0.15	**3.217** **
		(-7.716)	(2.755)	(3.008)	(3.652)	(2.330)			(13.076)	(-1.424)	(-0.632)	(0.419)	(0.713)		
	01/07-12/07	0.002	-0.032	**-0.047** **	-0.028	-0.017	0.09	1.630	**0.139*** **	0.033	0.035	0.007	**-0.084** **	0.31	**7.794*** **
		(0.117)	(-1.371)	(-2.076)	(-1.222)	(-0.622)			(5.541)	(1.205)	(1.298)	(0.266)	(-2.598)		
	01/08-12/08	**0.096** **	0.015	-0.010	-0.011	-0.029	0.03	0.561	**-0.392*** **	-0.030	-0.034	-0.043	-0.053	0.04	0.682
		(2.606)	(0.370)	(-0.246)	(-0.281)	(-0.612)			(-14.190)	(-0.985)	(-1.172)	(-1.395)	(-1.472)		
	01/09-12/09	**0.149*** **	-0.057	-0.049	-0.034	-0.009	0.03	0.669	**0.352*** **	-0.022	**-0.075***	**-0.078***	0.005	0.11	**2.555** **
		(3.263)	(-1.145)	(-1.012)	(-0.676)	(-0.162)			(8.487)	(-0.484)	(-1.678)	(-1.684)	(0.092)		
	01/10-12/10	-0.043	0.008	0.009	0.044	0.069	0.07	1.432	**0.170*** **	0.001	0.029	0.051	**0.080***	0.10	**2.240***
		(-1.111)	(0.196)	(0.219)	(1.010)	(1.412)			(5.118)	(0.016)	(0.822)	(1.380)	(1.855)		
3-Jahres	01/03-12/05	**-0.039*** **	-0.018	-0.010	-0.019	-0.019	0.05	0.693	**0.219*** **	-0.001	0.003	0.019	0.023	0.09	1.358
Zeitraum		(-2.726)	(-1.154)	(-0.633)	(-1.214)	(-1.034)			(13.104)	(-0.040)	(0.150)	(1.025)	(1.071)		
	01/04-12/06	**-0.083*** **	0.003	0.001	0.007	-0.010	0.03	0.455	**0.177*** **	-0.001	0.002	0.013	0.036	0.07	1.154
		(-5.905)	(0.183)	(0.069)	(0.458)	(-0.539)			(9.329)	(-0.029)	(0.091)	(0.611)	(1.433)		
	01/05-12/07	**-0.113*** **	**0.041*** **	**0.044*** **	**0.049*** **	**0.023***	0.32	**6.615*** **	**0.205*** **	0.009	0.015	0.010	-0.009	0.07	1.198
		(-11.197)	(3.762)	(4.146)	(4.476)	(1.775)			(15.455)	(0.612)	(1.090)	(0.688)	(-0.509)		
	01/06-12/08	**0.036** **	-0.013	-0.021	-0.002	-0.009	0.05	0.841	**-0.131*** **	**0.062*** **	**0.071*** **	**0.065*** **	**0.051*** **	0.24	**5.730*** **
		(2.025)	(-0.652)	(-1.088)	(-0.079)	(-0.410)			(-9.127)	(3.922)	(4.672)	(4.067)	(2.752)		
	01/07-12/09	**0.051** **	0.018	-0.008	0.009	0.010	0.07	1.292	**-0.081*** **	**0.036** **	**0.034** **	**0.026***	**0.032***	0.10	1.905
		(2.544)	(0.826)	(-0.355)	(0.406)	(0.393)			(-6.297)	(2.511)	(2.510)	(1.800)	(1.911)		
	01/08-12/10	0.009	0.036	0.022	0.032	**0.067** **	0.11	**2.140***	**-0.089*** **	0.035	**0.041***	**0.057** **	**0.107*** **	0.23	**5.347*** **
		(0.468)	(1.606)	(1.029)	(1.413)	(2.544)			(-4.335)	(1.547)	(1.882)	(2.504)	(4.034)		
5-Jahres	01/03-12/07	**-0.046*** **	-0.011	-0.008	-0.009	-0.016	0.04	0.539	**0.188*** **	0.019	**0.024***	**0.027***	**0.031***	0.08	1.135
Zeitraum		(-5.066)	(-1.142)	(-0.844)	(-0.873)	(-1.335)			(14.280)	(1.288)	(1.691)	(1.845)	(1.832)		
	01/04-12/08	0.000	**-0.025***	-0.017	-0.010	**-0.034** **	0.14	**2.189***	**-0.028** **	**0.049*** **	**0.049*** **	**0.048*** **	**0.060*** **	0.25	**5.114*** **
		(0.036)	(-1.858)	(-1.261)	(-0.702)	(-2.146)			(-2.507)	(3.949)	(4.163)	(3.857)	(4.091)		
	01/05-12/09	**0.050*** **	**-0.041*** **	**-0.035** **	**-0.044*** **	**-0.047** **	0.14	**2.384** **	**0.043*** **	0.014	0.019	0.012	0.022	0.05	0.964
		(3.544)	(-2.679)	(-2.320)	(-2.863)	(-2.617)			(4.051)	(1.168)	(1.635)	(0.988)	(1.592)		
	01/06-12/10	**0.032** **	-0.017	**-0.026***	-0.017	-0.014	0.06	0.926	0.009	**0.033** **	**0.041*** **	**0.034** **	**0.059*** **	0.19	**4.246*** **
		(2.223)	(-1.058)	(-1.695)	(-1.063)	(-0.746)			(0.766)	(2.536)	(3.253)	(2.572)	(3.909)		

Alternative Fondsbewertung mithilfe von Performancemaßen: 4-Faktor-Alpha und geometrisches Mittel (in-sample: 5 Jahre). *, ** und *** zeigen die Signifikanz auf dem 10%, 5% und 1% Niveau an.

Part IV:
Ratings and Performance of German Mutual Funds – A Comparison of Feri Trust, Finanztest, and FondsNote

Abstract

This study addresses the question whether the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote are able to predict the future performance of German equity mutual funds. Moreover, this study analyzes whether predictability is improved significantly when combining the three fund ratings. The reason is that investors compare fund ratings before decision-making. They invest in funds which uniformly have the best fund rating from different rating agencies.

It is shown that predictability is highest for the Feri Trust rating and the FondsNote compared to the Finanztest-Bewertung. Nevertheless, the three fund ratings can hardly predict future fund performance. Predictability is enhanced when combining all three fund ratings. However, it depends on the particular fund rating combination, the chosen performance measure and the post-rating period.

JEL Classification: G1, G11, G12, G14

Keywords: fund rating, ranking, rating persistence, performance persistence, predictability

I would like to thank Sigrid Müller and Stefan Schöne from the Institute of Finance and Sigbert Klinke from the Institute of Statistics for valuable comments and suggestions. Feri Trust, Stiftung Warentest and Axel Springer/FondsConsult provided me with their monthly fund ratings.

1 Introduction

Mutual funds are rated by several rating agencies that help in reducing extensive information to one quality label. Highly rated funds are especially advertised since investors combine these fund ratings with high performance expectations. Consequently, the grading of mutual funds seems to influence the behavior of investors. In this context, the question arises whether fund ratings are indeed able to predict the future performance of examined funds and whether investors should, therefore, include fund ratings in their decision-making.

Recent studies only evaluate the predictability of each fund rating on its own, although investors rather compare fund ratings before decision-making. They invest in funds which have the best fund rating from different rating agencies at the same time, respectively. This study, therefore, analyzes whether the predictability is improved significantly when comparing different fund ratings. To my knowledge, this is the first study to consider the predictability of fund rating combinations.

This study evaluates the predictability of three fund ratings: the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote. Recent studies mainly evaluate the predictability of the Morningstar rating. A great body of literature exists especially for the U.S. market. Most studies find the Morningstar rating to be a rather poor predictor of future fund performance (e.g. Blake and Morey (2000), Kräussl and Sandelowsky (2007), Garnier and Pujol (2007), Duret et al. (2008), Antypas et al. (2009), Füss et al. (2010), and Hereil et al. (2010)).

Few studies examine ratings from further fund rating agencies. Müller and Weber (2014), for instance, analyze the Finanztest-Bewertung of Stiftung Warentest, a well-known consumer protection agency in Germany. Their results suggest predictability of future fund performance, contrary to the studies analyzing the Morningstar rating. Meinhardt (2014) extends the analysis of the predictability to the Feri Trust rating and the FondsNote in addition to the Finanztest-Bewertung. The results indicate that the FondsNote has the highest forecasting ability compared to the other two fund ratings.

The predictability of the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote is analyzed with two different methods in this study. First, the three fund rating agencies are analyzed by the stability of their ratings over time. The idea behind this approach is that changes in ratings should be rare if the rating agency is able to predict the future performance of funds. The longer the period of time is until rating changes can be identified, the higher the forecasting ability of the fund rating agency (e.g. Garnier and Pujol (2007), Duret et al. (2008), and Hereil et al. (2010)).

Second, this study analyzes the post-rating performance of examined funds. Compared to the existing literature, fund ratings are considered each month within the examination period. The advantage is that all rating changes are included and the results are independent of the current market phase. Recent studies only evaluate the predictability at specific points in time within the examination period. For example, Blake and Morey (2000) examine the Morningstar rating in January each year from 1992 to 1997. Kräussl and Sandelowsky (2007), Füss et al. (2010), and Meinhardt (2014) analyze fund ratings in March, April, and December, respectively.

This study is divided into five sections. Preliminary considerations about the rating methodology used by the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote and their implications for investor decisions are discussed in Section 2. Section 3 presents the dataset and the methodology used in this study. Section 4 discusses the results and Section 5 presents the conclusion.

2 Fund Ratings versus Fund Rankings

Fund rating agencies have different methodologies to evaluate funds. Strictly speaking, the grading of a fund is based on a rating or on a ranking, albeit combinations between both methods are also possible. Within fund ratings, funds are compared among each other with respect to risk and return. Quantitative factors such as the performance and the risk of loss are analyzed. Moreover, the valuation process includes qualitative factors, which are used to assess the future fund performance. Qualitative factors are, for instance, an assessment of both the fund management quality and the investment process. Fund rankings have to be distinguished from fund ratings because the former only consider quantitative factors. Consequently, funds are sorted into ranking lists based on varying backward-looking factors. The fund's future performance is therefore not analyzed. For further differences and commonalities between both fund ratings and fund rankings see Meinhardt (2009).

This study considers three fund ratings that evaluate funds differently. First, the Feri Trust rating, first published in early 1999, is one of the most important fund ratings in Europe. Funds with Feri Trust rating are always evaluated quantitatively if the fund is registered more than five years for public sale. If the history of the fund is less than five years, qualitative factors are included in the valuation process. Such Feri Trust ratings are only created on behalf of the investment company and have no significance for this study. Consequently, the Feri Trust rating can be classified as a fund ranking.

Second, the Finanztest-Bewertung, first published by Stiftung Warentest in 1991, is examined in this study. Stiftung Warentest is a renowned German consumer protection organization and its fund ratings have a high reputation among investors. Just as the Feri Trust rating, the Finanztest-Bewertung does not take forecasts into account. The valuation process is exclusively based on backward-looking data and the Finanztest-Bewertung can therefore be classified as a fund ranking.

Third, this study examines the FondsNote which was jointly introduced by the Axel Springer Finanzen Verlag and FondsConsult in late 2002. FondsNote quantitatively allocates funds to one of five possible rating classes with a "1" for the best and a "5" for the most poorly rated funds. The final rating of the FondsNote can include quantitative as well as qualitative factors. Funds from the rating class "1" to "3" are always evaluated quantitatively and qualitatively. Lower rated funds receive a purely quantitative grade. Thus, the FondsNote does not only consist of historical data in comparison to the Feri Trust rating and the Finanztest-Bewertung. Consequently, the FondsNote is a mixture of a fund ranking and a fund rating. Table 1 presents further differences between the three fund rating agencies and their valuation processes. Meinhardt (2014) also analyzes these three fund rating agencies and presents a more detailed analysis of the underlying methodology.

Table 1: Characteristics of Fund Ratings

	Feri Trust rating	Finanztest	FondsNote
Number of quantitative factors	12	2	3
Qualitative factors	no	no	yes
Categories by investment style	none	none	none
Best rating*	(A)	substantially above average	1
Worst rating*	(E)	substantially below average	5
Proportion of funds with best rating	none	none	none

* Finanztest names the best-rated funds as "stark überdurchschnittlich" in German. On the contrary, the worst-rated funds are named as "stark unterdurchschnittlich".

How do the different valuation methods affect the forecasting ability of fund ratings and fund rankings? Is it important to know before investor's decision whether the valuation process depends on quantitative factors only or on both quantitative and qualitative factors? Theory implies that the predictability depends on whether the valuation process only relies on past information or also includes qualitative factors that attempt to assess the fund's future performance. Consequently, the Feri Trust rating and the Finanztest-Bewertung should have a lower forecasting ability than the FondsNote, which evaluates funds quantitatively and qualitatively. If fund ratings do not have a higher predictability than fund rankings, one can ask why rating agencies evaluate funds qualitatively at all. The assessment of the fund management and the evaluation of the in-

vestment process, for instance, induce additional costs, which can only be compensated for by higher license fees for advertising. This would result in a competitive disadvantage compared to fund rankings that only rely on quantitative factors. By contrast, if fund ratings can successfully separate between well and poorly performing funds compared to fund rankings, these rating agencies should enhance their transparency of methodology. Consequently, the market acceptance would increase because investors could better differentiate between rating agencies with high and low predictability.

The forecasting ability of fund ratings can be analyzed by two different approaches: the rating persistence and the performance persistence. For the sake of simplicity, fund ratings mean ratings as well as rankings from here on. A fund rating persists over time if a fund is assigned to a rating that stays stable. A rated fund, given the highest rating available, should perform better in the future than worse rated funds. Consequently, rating persistence means that the fund rating may not change over time. Moreover, a fund rating has a forecasting ability if the performance of rated funds persists over time. Performance persistence can be noticed, for example, if the best-rated funds have a higher future performance than worse rated funds.

3 Data and Methodology

3.1 Mutual Fund Sample

Analyses of the predictive power are based on the monthly published ratings of equity mutual funds, with investment focus on Germany, of the Feri Trust rating (FTR), the Finanztest-Bewertung (FT), and the FondsNote (FN). Fund ratings have been made available directly from the rating agencies for the period from December 2002 to December 2009 and include surviving as well as merged or liquidated funds. Hence, the dataset is free of survivorship bias. Necessary data for the analysis such as monthly fund prices (total return index) are available from Bloomberg. Fund prices are checked with data from Thomson Reuters Datastream to eliminate errors and gaps. Fama-French factors for Germany are gathered from the study of Brückner et al. (2014).

Overall, 105 funds have received a Feri Trust rating in the category “German Equities” within the examination period. 139 and 96 funds have got a Finanztest-Bewertung and a FondsNote, respectively. However, the rating discontinued for 46 funds with Feri Trust rating, for 77 funds with Finanztest-Bewertung and for 17 funds with FondsNote due to liquidation, mergers or changes in investment focus. Funds rated average and below were particularly involved. Table 2 presents an overview of examined funds per fund rating across time.

Table 2: Descriptive Statistics – Fund Universe

	Feri Trust Rating	Finanztest	FondsNote
12/2002	81	89	58
12/2003	75	80	65
12/2004	71	76	75
12/2005	70	80	78
12/2006	68	81	78
12/2007	69	80	80
12/2008	66	85	92
12/2009	59	75	90

The table presents the number of funds with a Feri Trust rating, FondsNote and Finanztest-Bewertung within the examination period. This includes all equity mutual funds with investment focus on "German Equities".

The literature has corrected for the resulting survivorship bias in several ways. Both Blake and Morey (2000) and Kräussl and Sandelowski (2007) assume a balanced reinvestment in remaining funds after fund liquidations or mergers. Moreover, Morey and Gottesman (2006) use two alternative procedures to avoid a survivorship bias. The first method is based on the fact that those funds are eliminated from the dataset that have out-of-sample returns of less than 12 months. The second method assumes a reinvestment in a very similar fund. Criteria are, for instance, the latest fund rating, the total expense ratio, the turnover and the assets under management.

In this study, returns of liquidated funds are proceeded with returns of very similar funds according to the second method of Morey and Gottesman (2006). As opposed to Morey and Gottesman (2006), two decision criteria are used for this purpose because the availability of data is scarce for German funds compared to U.S. funds. First, both fund ratings must be identical at the time of liquidation or merger. Second, these two funds should have similar returns in the previous 12 months. If a fund does not receive a fund rating anymore due to a merger, I will pursue with returns of the remaining fund from the date of the merger. This requires the condition that the continued fund has its focus on "German Equities".

3.2 Testing for Predictive Abilities

Rating persistence

The forecasting ability of different fund ratings can be determined by the persistence of the rating over time. This means that the predictability of a fund rating is high if the fund rating does not change in the short-term. The predictability of the rating is low if there are quick changes in the fund rating. Consequently, a fund rating, which includes the future development of a fund in their valuation process, should not change as quickly as a fund rating that only depends on past information. This study determines the rating persistence of fund ratings based on homogeneous

Markov chain modeling. This is in line with previous research (e.g. Garnier and Pujol (2007), Duret et al. (2008), and Hereil et al. (2010)).

In this context, a fund rating system is defined as a set of different valuation classes $R = \{R_1, R_2, \dots, R_K\}$. All examined fund ratings use five valuation classes to distinguish between the best and the worst performing funds. Moreover, this study introduces the class "NR" for unrated funds in order to offset a potential bias. Unrated funds are funds that are included in the universe of rated funds or excluded from the universe due to liquidation, mergers or changes in investment focus. Consequently, analyzed funds can be sorted monthly into six different valuation classes.

Assuming that fund i has the fund rating R_k at time t, $\pi_{(i)}(s, k_1; t, k_2)$ is defined as the probability that the valuation class of fund i changes from R_{k_1} at time s to R_{k_2} at time t. Thus:

$$\pi_{(i)}(s, k_1; t, k_2) = \Pr\{R_{(i)}(t) = R_{k_2} | R_{(i)}(s) = R_{k_1}\}. \tag{1}$$

This dynamic system can be modeled in the context of a Markov chain setting. Assuming that $\pi(t)$ is a homogeneous Markov matrix and represents the transition matrix between $t = 0$ and t. Then there exists a Markov generator Λ, so that:

$$\pi(t) = \exp(t\Lambda). \tag{2}$$

Using the method of maximum likelihood, the monthly transition matrix $\hat{\pi}(t)$ can be estimated for each fund rating. Monthly matrices for the three fund ratings are presented in Table A1 in the Appendix. The Markov generator Λ can be estimated with $\widehat{\Lambda}$, which is defined by:

$$\widehat{\Lambda} = \frac{1}{t} \ln \hat{\pi}(t). \tag{3}$$

Thus:

$$\tilde{\pi}(t) = \exp\left(t\widehat{\Lambda}\right). \tag{4}$$

This Markov generator $\widehat{\Lambda}$ is only valid if each off-diagonal element is non-negative and all its row-sums are zero. Negative off-diagonal elements lead to a generator $\widehat{\Lambda}$ that is not Markov and, therefore, $\tilde{\pi}(t)$ is not a Markov transition matrix. Israel et al. (2001) present two methods to adjust the generator in order to obtain a valid Markov generator. This study uses the second approach (IRW-2).

Dummy variable regression

Instead of assessing the rating persistence of fund ratings, the forecasting ability can also be analyzed by the performance persistence of rated funds over time. Fund ratings are mainly based on the past performance of funds. Consequently, well performing funds in the past get a better rating than low performing funds. Performance persistence in the context of fund ratings, therefore, means that the best-rated funds should have a higher future performance than worse rated funds.

This study determines the forecasting ability of the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote by analyzing the average post-rating performance of an investment in funds of one rating class. Funds are assigned to quintiles based on their fund ratings in month t. To compare the average performance of formed quintiles, previous studies apply a cross-sectional dummy variable regression, which was first introduced by Blake and Morey (2000). The baseline dummy variable regression is of the form:

$$S_i = \beta_1 + \beta_2 D2_i + \beta_3 D3_i + \beta_4 D4_i + \beta_5 D5_i + \varepsilon_i, \tag{5}$$

where S_i is the post-rating performance of fund *i*, and $D2_i$ through $D5_i$ are dummy variables that refer to the fund rating of fund *i*. For example, $D5_i$ is 1 if fund *i* receives the best rating and 0 otherwise. Hence, the group of the worst-rated funds forms the reference class, and the performance of all other quintiles is measured in relation to them. Fund ratings can predict future fund performance if the coefficients β_2 to β_5 are significantly positive and increase in the form that $\beta_2 < \beta_3 < \beta_4 < \beta_5$.

Blake and Morey (2000) examine the post-rating performance of funds with Morningstar rating in January of each year during the observation period. Kräussl and Sandelowsky (2007), Füss et al. (2010), and Meinhardt (2014) consider the rating of the funds in March, April, and December, respectively. Compared to previous studies, I consider the fund ratings each month within the examination period via multiple cross-sectional regressions, using the methodology of Fama and MacBeth (1973). Hence, a bias that only arises due to the market period in which an investment is made is circumvented.

The post-rating performance of all funds analyzed is measured for one- and three-year periods and based on the Carhart four-factor alpha and the geometric mean of monthly returns. Using multiple cross-sectional regressions for each month during the observation period leads to overlapping return periods. For example, the one-year post-rating performance of two consecutive months overlaps in eleven of twelve months. The resulting serial correlation in regression residuals is corrected with the Newey-West procedure and a lag length of eleven months.

Just as Müller and Weber (2014), this study applies the monotonicity test of Patton and Timmermann (2010) in order to analyze whether the coefficients β_2 to β_5 of the dummy variable regression are indeed strictly increasing. Furthermore, a one-factorial ANOVA is used to detect all significant differences between classes. The dummy variable regression only compares classes against a reference class and as a consequence, significant differences between other classes cannot be detected, although the F-statistic is highly significant. It is important to notice that the one-factorial ANOVA does not correct for the serial correlation induced by overlapping performance periods. The t-statistics only represent an upper bound of significant class differences. To avoid an alpha error accumulation, I use Tukey's (HSD) or Games-Howel's post-hoc test correction depending on whether heteroscedasticity between classes is detected or not.

Fund ratings, fund rating combinations, and alternative predictors

This study examines the rating persistence and performance persistence for four different fund samples to get an overall insight into the forecasting ability of the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote. First, the three fund ratings are analyzed separately. In this way, I can answer the question whether the examined fund ratings can predict the future fund performance on their own. Moreover, this study can verify whether the forecasting ability significantly differs between fund ratings that rely on quantitative factors only and fund ratings with quantitative and qualitative factors.

Second, the question whether fund rating predictability is improved significantly when comparing these three fund ratings is taken into account. For this purpose, four different fund rating combinations are examined resulting in three pairwise comparisons and one connected analysis of the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote. It is important to notice that fund ratings are always equally-weighted in all combinations.

Third, this study compares the forecasting ability of the fund rating quintiles to two alternative predictors (the four-factor alpha and the geometric mean of monthly returns). In this way, I can analyze whether fund ratings like the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote give additional information to investors compared to performance indicators. If fund ratings cannot better differentiate between high and low performing funds, the existence of fund rating agencies can be questioned. To analyze this issue, funds are sorted into quintiles based on the pre-rating performance of these two alternative predictors. The performance is measured for 60 month periods because both the Feri Trust rating and the Finanztest-Bewertung rely on this evaluation period. The results barely change using a four-year performance window as done by the FondsNote.

Fourth, the fund ratings, which are formed with the alternative predictors, are compared to each other just as the original fund ratings.

4 Results

4.1 Rating Persistence

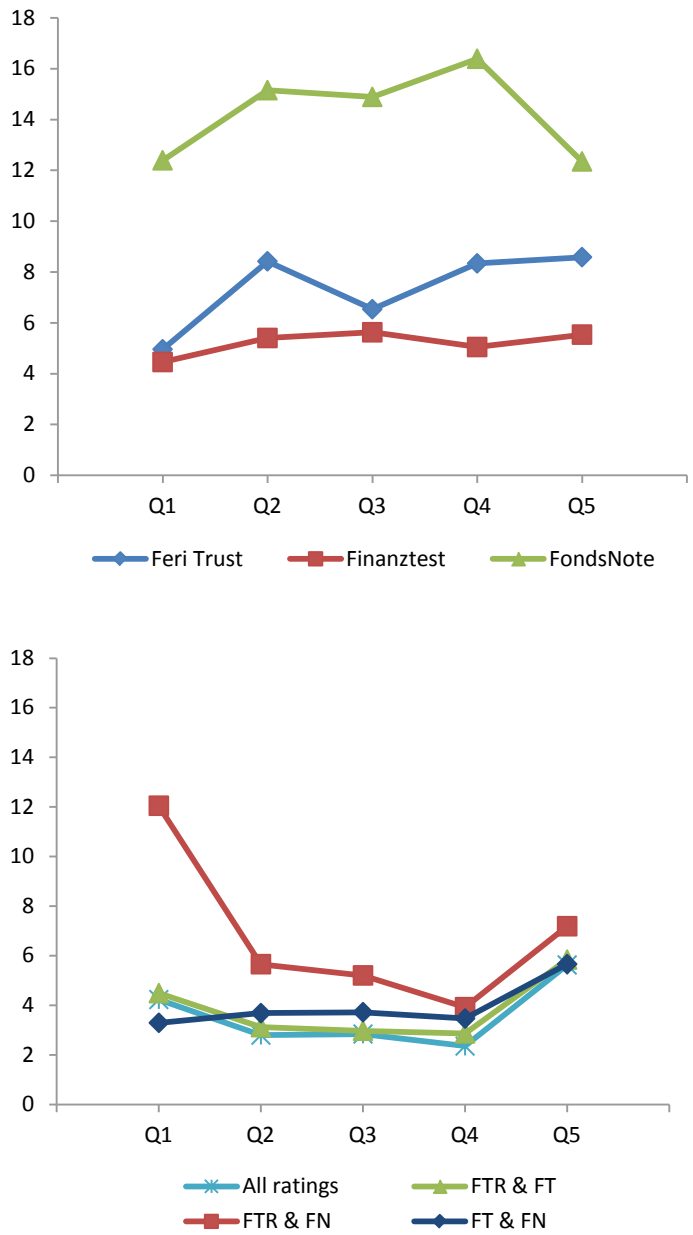
Original fund ratings

This section provides the results based on the transition matrices estimated for the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote. Using a valid Markov generator, transition matrices can be estimated for different periods of time t. For the sake of brevity, this study does not present transition matrices depending on time but rather shows the period of time until rated funds change their rating with a probability of 50%. This measure can be viewed as a rating persistence measurement.

Figure 1 indicates that the valuation process of the rating agency matters. Fund ratings that include both quantitative and qualitative factors show the best results. Consequently, the rating persistence is highest for the FondsNote compared to the Feri Trust rating and the Finanztest-Bewertung. Funds rated with the Finanztest-Bewertung have the quickest changes in their ratings. The period of time until rated funds change their rating amounts to 12.35, 16.39, 14.89, 15.16, and 12.40 months for the best to the worst FondsNote (Q5 to Q1). This is approximately two times higher than the rating persistence measured with the Feri Trust rating (8.59, 8.34, 6.54, 8.42, and 4.96) and three times higher than with the Finanztest-Bewertung (5.54, 5.05, 5.63, 5.40, and 4.46). Nevertheless, fund ratings of all three rating agencies only stay stable for a short period of time (see also Table A2.1 to Table A2.3 in the Appendix).

The combination of different fund ratings does not improve the rating persistence. This is true for combinations of all three fund ratings. Figure 1 shows the results for all fund rating combinations. Combining the Feri Trust rating and the FondsNote delivers the highest rating persistence compared to all other fund rating combinations. That is not surprising since these two fund ratings, considered independently, yield the highest rating persistence. Moreover, the period of time is frequently the highest for the extreme ratings, which are defined as the best and the worst rating quintiles, but in sum also low. Consequently, an investment strategy that only consists of the combination of the three fund ratings does not lead to a higher rating persistence.

Figure 1: Rating Persistence – Fund Ratings and Fund Rating Combinations

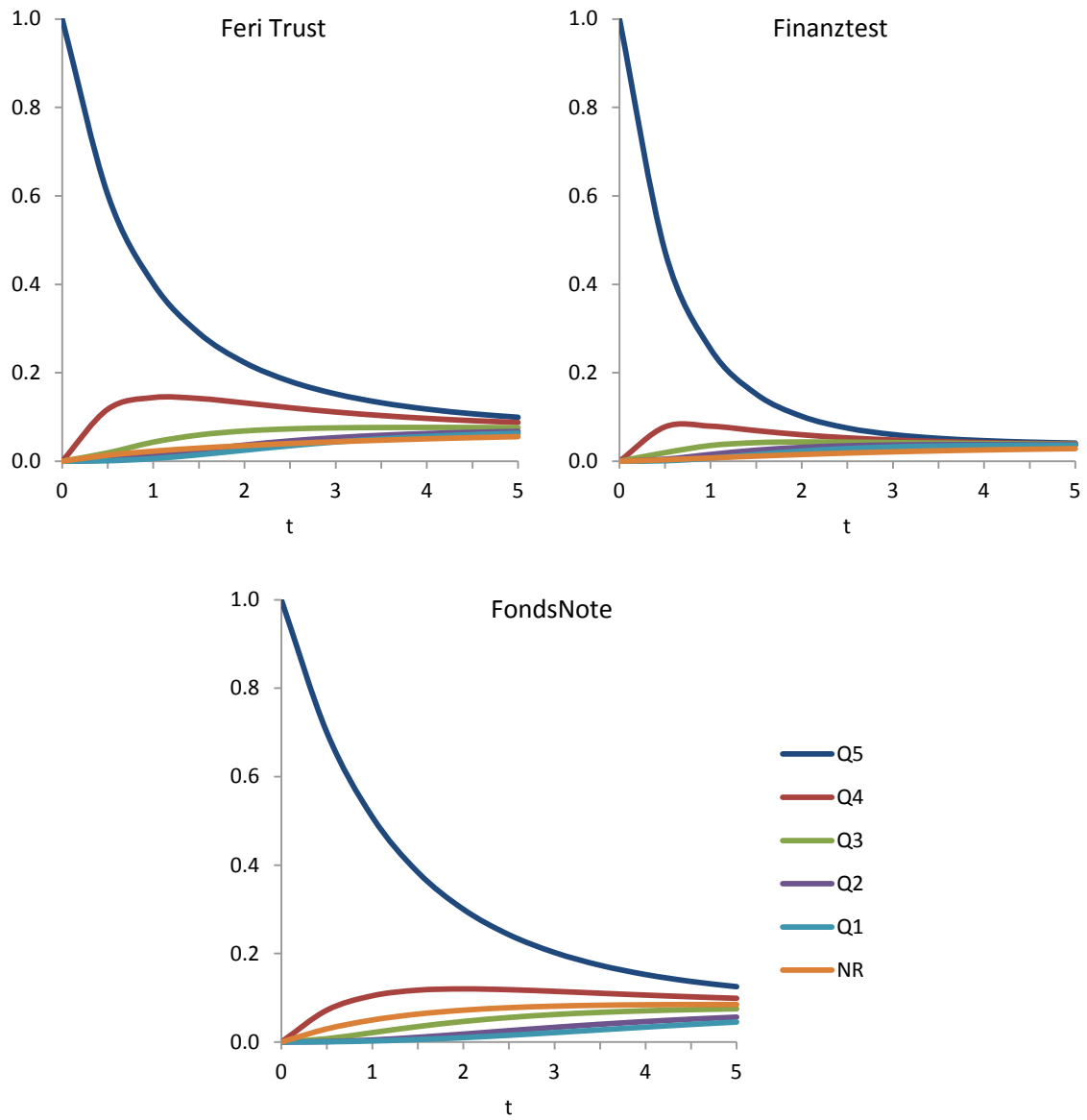


The figure presents the period of time (in months) until rated funds change their rating with a probability of 50%. The period of time is separately measured for different rating classes. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0. Fund rating combinations are based on the Feri Trust rating (FTR), the Finanztest-Bewertung (FT), and the FondsNote (FN).

Figure 2 takes a closer look into the development of rating changes of the best-rated funds over time. The second-best-rated to worst-rated funds present similar results. Figure 2 shows the probability that a fund gets the best rating at time t depending on the rating at time 0. The results

confirm that the rating persistence is highest for the FondsNote, which can be seen in the lower slope of the graph (quintile 5) compared to the Feri Trust rating and the Finanztest-Bewertung.

Figure 2: Transition Probabilities for the Best-Rated Funds



The figure presents the probabilities that a fund is rated by the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote at time 0 and gets the best fund rating at time t (in years). The best-rated to the worst-rated funds are represented by quintile 5 (Q5) to 1 (Q1). "NR" states funds that are rated at time 0 but are not rated anymore at time t.

Interestingly, Figure 2 also indicates the existence of a long-run equilibrium with regard to the probability of rating changes. The probabilities have a fixed distribution, which will be achieved after approximately five years, irrespective of today's fund rating. For example, the probability for a fund to get the best Feri Trust rating is 7% after five years. This probability amounts to 3% for funds with Finanztest-Bewertung and 8% for funds that are rated with a FondsNote today. The

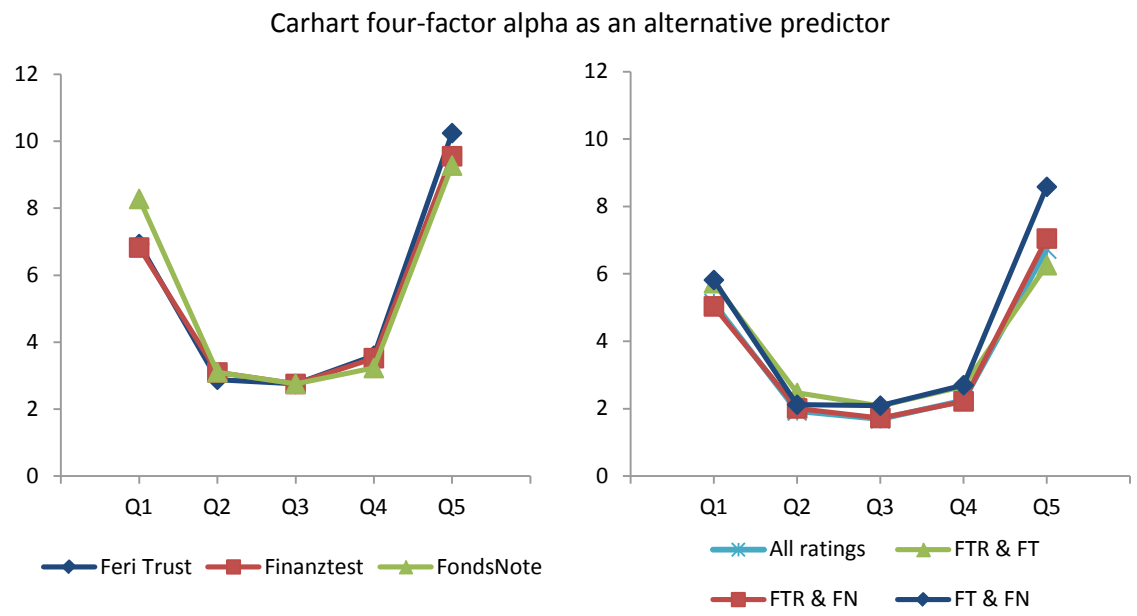
distribution of probabilities for the second-best-rated to the worst-rated funds amounts to 16%, 20%, 19%, and 5% for the Feri Trust rating and is similar to the Finanztest-Bewertung. The Feri Trust rating is discontinued for 33% of all the examined funds after five years. Only the FondsNote yield slightly different results. It has to be emphasized that funds achieve the third-best rating with a probability of 30%, which is 10% higher than with the Feri Trust rating. Moreover, only 7% of all examined funds will not be rated anymore after five years.

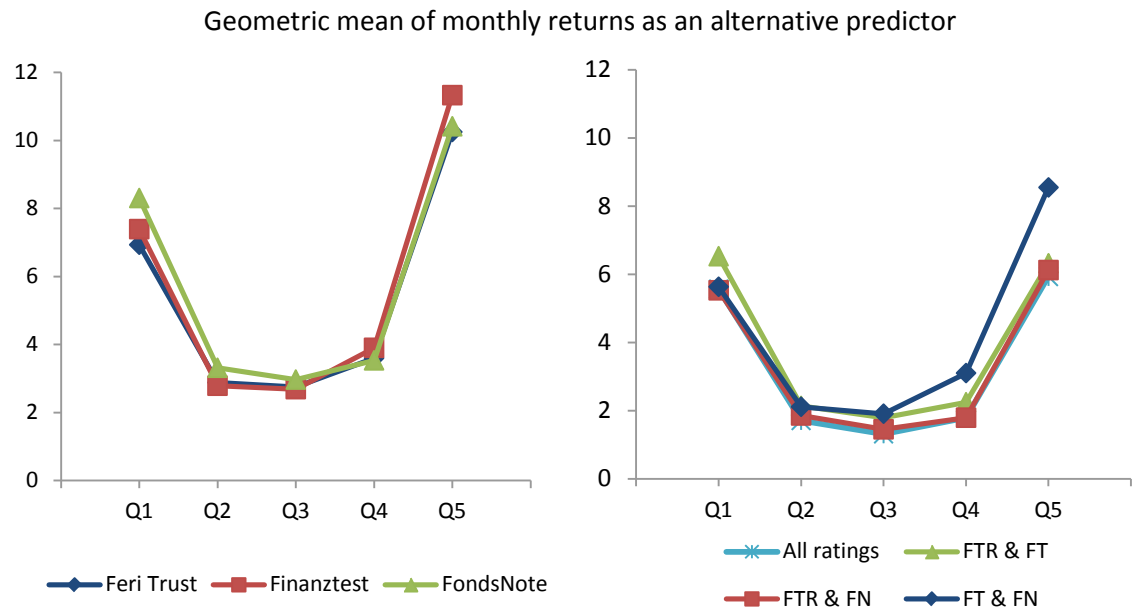
Alternative predictors

Alternative ratings based on the Carhart four-factor alpha and the geometric mean of monthly returns yield slightly different results (Figure 3). First, extreme rating classes have a higher rating persistence than measured with both the Feri Trust rating and the Finanztest-Bewertung on their own. This is not true for funds that get the best and the worst FondsNote. Funds in quintile 2 to 4, which are rated with the alternative predictors, have a substantially lower rating persistence than with the original fund ratings.

The combination of fund ratings based on alternative predictors does not significantly increase the rating persistence. The results are very similar to the rating persistence measured for alternative predictors on their own. The resulting U-form from quintile 1 to 5 is in line with the results for the S&P fund rating which are presented in the study of Duret et al. (2008). Instead, Hereil et al. (2010) find for the Morningstar rating that funds with an average rating have the highest rating persistence.

Figure 3: Rating Persistence – Alternative Predictors





The figure presents the period of time (in months) until rated funds change their rating with a probability of 50%. The period of time is separately measured for different rating classes. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0. Fund rating combinations are based on the Feri Trust rating (FTR), the Finanztest-Bewertung (FT), and the FondsNote (FN).

4.2 Performance Persistence

Original fund ratings

Can we conclude from a low rating persistence over time that the ability to predict the post-rating performance of funds is also low? This section answers the question whether the performance persistence is in line with the analyzed rating persistence, which prefers the FondsNote to the Feri Trust rating and the Finanztest-Bewertung. Regression coefficients of the dummy variable analysis are presented for all three fund ratings depending on one- and three-year post-rating performances, respectively.

Table 3.1 points out that the forecasting ability of examined fund ratings is poor, irrespective of the performance measure. The Feri Trust rating, the Finanztest-Bewertung as well as the FondsNote can scarcely differentiate between future high and low performing funds. This is in contrast to the results of the rating persistence measurement. The difference in the average post-rating performance between the best and the worst-rated funds is frequently positive but it is rarely significantly different from zero. Significant performance differences can only be seen for the Feri Trust rating and the FondsNote based on the four-factor alpha and a three-year observation period. Moreover, the monotonic relation tests (MR^{All}) in the last column essentially do not reject the null hypothesis. Consequently, the performance of the quintiles from 1 to 5 mainly does not increase monotonically.

Interestingly, Panel A shows that higher rated funds (quintile 2 to 5) are not able to outperform their indices. For example, the performance difference of funds with Feri Trust rating being assigned to quintile 5 is 3.8% per year for a three-year period. That still implies a negative four-factor alpha of -1.5% providing that the performance of funds belonging to quintile 1, the reference class, is -5.3%. Compared to the four-factor alpha, in Panel B, the absolute performance of all quintiles is positive. One reason is that the geometric mean used in this study measures the raw return of fund quintiles without any risk adjustments.

Table 3.1: Dummy Variable Regression – Original Fund Ratings

Fund rating		β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.	MR^{All}
				Panel A: Carhart four-factor alpha					
Feri Trust	1 year	**-0.053*** **	0.022	0.024	0.026	0.025	0.021	**2.219***	0.129
		(-4.110)	(0.949)	(0.819)	(0.922)	(0.884)			
	3 years	**-0.053*** **	0.002	**0.030***	**0.036***	**0.038** **	0.064	**7.227*** **	0.007
		(-4.359)	(0.084)	(1.679)	(1.946)	(2.271)			
Finanztest	1 year	-0.022	-0.004	-0.004	0.001	-0.023	0.017	1.840	0.893
		(-0.569)	(-0.098)	(-0.101)	(0.013)	(-0.387)			
	3 years	-0.018	-0.001	0.005	0.009	0.000	0.005	0.498	0.583
		(-2.465)	(-0.065)	(0.336)	(0.500)	(-0.026)			
FondsNote	1 year	**-0.043***	0.030	0.020	0.016	0.020	0.015	1.634	0.986
		(-0.808)	(0.493)	(0.329)	(0.279)	(0.324)			
	3 years	**-0.042*** **	0.027	**0.028***	0.032	**0.044*** **	0.059	**6.530*** **	0.003
		(-3.500)	(1.010)	(1.704)	(1.647)	(2.601)			
				Panel B: Geometric mean of monthly returns					
Feri Trust	1 year	0.113	0.005	0.006	0.008	0.014	0.000	0.035	0.016
		(1.606)	(0.061)	(0.063)	(0.080)	(0.127)			
	3 years	**0.065*** **	-0.010	0.009	0.010	0.014	0.004	0.466	0.338
		(3.387)	(-0.292)	(0.290)	(0.284)	(0.469)			
Finanztest	1 year	**0.126***	0.000	0.004	0.006	-0.031	0.003	0.313	0.584
		(1.822)	(-0.005)	(0.038)	(0.057)	(-0.301)			
	3 years	**0.074*** **	0.003	0.011	0.012	-0.009	0.003	0.362	0.935
		(5.939)	(0.094)	(0.365)	(0.386)	(-0.324)			
FondsNote	1 year	**0.132***	-0.002	-0.004	-0.004	-0.004	0.000	0.005	0.247
		(1.820)	(-0.020)	(-0.044)	(-0.036)	(-0.037)			
	3 years	**0.068*** **	0.008	0.012	0.012	0.015	0.002	0.158	0.014
		(3.932)	(0.196)	(0.380)	(0.365)	(0.514)			

The table presents the results from the monthly regressions of fund performance on the rating quintiles for the following fund ratings: the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. t-statistics are adjusted using the Newey-West procedure with a lag of eleven months or thirty-five months depending on the post-rating period. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively. MR^{All} presents the p-value of the monotonic relation test of Patton and Timmermann (2010).

In order to get a better understanding of significant performance differences between all analyzed quintiles, Table 3.2 provides the results of the one-factorial ANOVA analysis. The results support the poor forecasting ability of examined fund ratings, but considering the four-factor al-

pha, the three-year out-of-sample performance of funds with Feri Trust rating in quintile 2 is always significantly lower than the performance of higher rated funds. Further significant differences in performance are in line with the dummy variable analysis mentioned earlier. Using the geometric mean as performance measure, differences in performance between quintiles cannot be detected and, therefore, are not presented in this study.

Table 3.2: ANOVA – Original Fund Ratings

		Feri Trust		Finanztest		FondsNote	
		1 year	3 years	1 year	3 years	1 year	3 years
(I)	(J)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)
				Carhart four-factor alpha			
β_1	β_2	0.022	0.002	-0.004	-0.001	0.030	**0.027***
	β_3	0.024	**0.030****	-0.004	0.005	0.020	**0.028****
	β_4	0.026	**0.036****	0.001	0.009	0.016	**0.032*****
	β_5	0.025	**0.038*****	-0.023	0.000	0.020	**0.044*****
β_2	β_3	0.002	**0.028****	0.000	0.006	-0.010	0.001
	β_4	0.005	**0.034*****	0.005	0.010	-0.014	0.005
	β_5	0.003	**0.036*****	-0.019	0.001	-0.010	0.017
β_3	β_4	0.003	0.006	0.005	0.004	-0.003	0.004
	β_5	0.002	0.008	-0.019	-0.005	0.000	0.016
β_4	β_5	-0.001	0.002	-0.024	-0.009	0.004	0.012

The table presents the post-hoc tests of the one-factorial ANOVA on the rating quintiles based on the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote, respectively. Depending on the results of the Levene-statistics, test statistics of post-hoc tests are either adjusted with Tukey (HSD) or Games-Howel. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

The poor forecasting ability of the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote determined in this study is in line with the existing literature. Although most previous studies have concentrated on the Morningstar rating, this study can scarcely find differences in performance predictability in comparison with the three analyzed fund ratings. The forecasting ability of the Finanztest-Bewertung is even lower than in the study of Müller and Weber (2014) regarding the German market. Meinhardt (2014), who analyzes the post-rating performance of fund ratings in December of each year, finds that the future fund performance is predicted better by the FondsNote compared to the Feri Trust rating and the Finanztest-Bewertung. Considering every month within the examination period, this study only finds a slightly better forecasting ability for the Feri Trust rating and the FondsNote compared to the Finanztest-Bewertung.

Combination of original fund ratings

The forecasting ability of fund ratings like the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote on their own is low. However, the post-rating performance of funds could be improved significantly when different fund ratings are compared with each other. The idea behind

this approach is that the three examined fund rating agencies evaluate their analyzed funds differently. When the fund rating assigned is still identical among the rating agencies, the validity of the rating should be higher compared to an isolated fund evaluation of one agency. This section outlines the predictability of combined fund ratings when funds receive the same rating in one month.

In Table 4.1, the results show that the forecasting ability can be enhanced when investment decisions do not only depend on one fund rating but include different fund ratings at the same time. It should be noted that not every fund rating combination increases the forecasting ability. Moreover, differences in the post-rating performance are primarily induced by the performance measure and performance period used.

In detail, the forecasting ability is enhanced using all three examined fund ratings as four-factor alpha performance predictor. Not only in this case, but also other fund rating combinations lead to significant four-factor alpha performance differences between analyzed quintiles. This result is confirmed by highly significant F-statistics in 7 out of 8 cases. However, the performances between quintiles do not differ from each other when using the geometric mean as performance measure regardless of the fund rating combination. Moreover, the MR-tests commonly do not reject the null hypothesis of increasing coefficients as we move from β_2 to β_5.

Table 4.1: Dummy Variable Regression – Combination of Original Fund Ratings

Fund rating		β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.	MR^{All}
				Panel A: Carhart four-factor alpha					
All ratings	1 year	**-0.087*** **	0.058	**0.059***	**0.064***	0.055	0.075	**8.535*** **	0.388
		(-3.108)	(1.607)	(1.663)	(1.654)	(1.405)			
	3 years	**-0.062*** **	0.004	**0.035***	**0.059*** **	**0.061** **	0.105	**12.259*** **	0.000
		(-4.671)	(0.156)	(1.802)	(3.106)	(2.359)			
FTR & FT	1 year	**-0.051*** **	0.018	0.024	0.018	0.011	0.011	1.140	0.533
		(-4.647)	(0.820)	(0.985)	(0.605)	(0.202)			
	3 years	**-0.037*** **	-0.022	0.011	0.018	0.019	0.044	**4.867*** **	0.384
		(-5.053)	(-1.021)	(0.882)	(1.121)	(1.035)			
FTR & FN	1 year	**-0.067*** **	0.049	0.041	0.037	**0.047** **	0.092	**10.632*** **	0.963
		(-16.788)	(1.446)	(1.561)	(1.338)	(2.138)			
	3 years	**-0.055*** **	**0.040** **	**0.027** **	**0.039*** **	**0.049*** **	0.108	**12.707*** **	0.553
		(-13.732)	(2.361)	(2.057)	(3.200)	(3.086)			
FT & FN	1 year	**-0.056*** **	0.045	0.029	0.031	0.005	0.056	**6.216*** **	0.967
		(-4.863)	(1.280)	(1.196)	(1.100)	(0.131)			
	3 years	**-0.046*** **	**0.038***	**0.022***	**0.038** **	0.030	0.057	**6.298*** **	0.884
		(-8.339)	(1.903)	(1.893)	(2.434)	(1.161)			

Panel B: Geometric mean of monthly returns

All ratings	1 year	0.079	0.045	0.047	0.058	0.044	0.007	0.706	0.298
		(1.002)	(0.423)	(0.430)	(0.504)	(0.414)			
	3 years	**0.057*** **	0.016	0.025	0.030	0.017	0.006	0.658	0.893
		(3.966)	(0.485)	(0.857)	(0.917)	(0.650)			
FTR & FT	1 year	**0.119***	0.003	0.006	-0.001	-0.023	0.002	0.181	0.831
		(1.757)	(0.034)	(0.057)	(-0.007)	(-0.236)			
	3 years	**0.073*** **	-0.001	0.010	0.002	-0.005	0.002	0.167	0.910
		(5.687)	(-0.021)	(0.369)	(0.056)	(-0.190)			
FTR & FN	1 year	**0.144*** **	-0.021	-0.016	-0.022	-0.037	0.003	0.291	0.827
		(5.220)	(-0.287)	(-0.205)	(-0.282)	(-0.445)			
	3 years	**0.062*** **	0.010	0.022	0.015	0.009	0.004	0.371	0.956
		(2.979)	(0.310)	(0.646)	(0.440)	(0.303)			
FT & FN	1 year	**0.105** **	0.026	0.026	0.029	0.004	0.003	0.282	0.327
		(2.500)	(0.309)	(0.288)	(0.311)	(0.050)			
	3 years	**0.052*** **	0.028	0.032	0.034	0.014	0.010	1.043	0.802
		(3.662)	(0.896)	(1.079)	(1.061)	(0.625)			

The table presents the results from the monthly regressions of fund performance on the rating quintiles for fund rating combinations based on the Feri Trust rating (FTR), the Finanztest-Bewertung (FT), and the FondsNote (FN). Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. t-statistics are adjusted using the Newey-West procedure with a lag of eleven months or thirty-five months depending on the post-rating period. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively. MR^{All} presents the p-value of the monotonic relation test of Patton and Timmerman (2010).

The results of the one-factorial ANOVA presented in Table 4.2 indicate that significant differences in performance between quintiles cannot only be seen for higher rated funds (quintile 2 to 5) compared to the worst-rated funds (quintile 1), but also between higher rated funds. Moreover, the post-rating performance of quintiles is often higher in absolute terms than the performance of quintiles formed due to the Feri Trust rating, the Finanztest-Bewertung, or the FondsNote, respectively.

Table 4.2: ANOVA – Combination of Original Fund Ratings

		All ratings		FTR & FT		FTR & FN		FT & FN	
		1 year	3 years	1 year	3 years	1 year	3 years	1 year	3 years
(I)	(J)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)
β_1	β_2	**0.058*** **	0.004	0.018	-0.022	**0.049*** **	**0.040*** **	**0.045*** **	**0.038*** **
	β_3	**0.059*** **	**0.035** **	0.024	0.011	**0.041*** **	**0.027*** **	**0.029*** **	**0.022** **
	β_4	**0.064*** **	**0.059*** **	0.018	0.018	**0.037*** **	**0.039*** **	**0.031*** **	**0.038*** **
	β_5	**0.055** **	**0.061*** **	0.011	0.019	**0.047*** **	**0.049*** **	0.005	**0.030** **
β_2	β_3	0.001	0.031	0.006	**0.033** **	-0.008	-0.013	-0.015	-0.017
	β_4	0.007	**0.055*** **	0.000	**0.040*** **	-0.012	-0.001	-0.013	0.000
	β_5	-0.003	**0.057*** **	-0.007	**0.041** **	-0.002	0.009	**-0.040** **	-0.008
β_3	β_4	0.005	**0.024** **	-0.006	0.007	-0.004	0.012	0.002	0.016
	β_5	-0.004	**0.026** **	-0.013	0.008	0.006	**0.022** **	-0.024	0.008
β_4	β_5	-0.010	0.002	-0.007	0.001	0.010	0.010	-0.026	-0.008

The table presents the post-hoc tests of the one-factorial ANOVA on the rating quintiles for fund rating combinations based on the Feri Trust rating (FTR), the Finanztest-Bewertung (FT), and the FondsNote (FN). The post-rating performance is measured with Carhart four-factor alpha. Depending on the results of the Levene-statistics, test statistics of post-hoc tests are either adjusted with Tukey (HSD) or Games-Howel. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

All in all, the rise in the forecasting ability depends on the fund rating combination, the performance measure and the post-rating period. Combining all three fund ratings, the predictability of future fund performance can be enhanced. Combinations of the Feri Trust rating and the FondsNote, which both predict in a similar way, only slightly increase the probability of significant performance differences between quintiles. The same is true when combining the Feri Trust rating with the Finanztest-Bewertung.

Alternative predictors

The forecasting ability is low for the three fund ratings examined on its own, but can partly be increased due to combinations of each other. The question is whether alternative predictors exist that can better predict the future fund performance. Alternative predictors are the Carhart four-factor alpha and the geometric mean of monthly returns.

Table 5.1: Dummy Variable Regression – Alternative Predictors

Fund rating		β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.	MR^{All}
Panel A: Carhart four-factor alpha as alternative predictor and performance measure									
Feri Trust	1 year	-0.031	0.010	0.004	0.004	0.003	0.003	0.342	0.804
		(-1.280)	(0.310)	(0.116)	(0.116)	(0.071)			
	3 years	**-0.050*** **	0.010	0.015	0.023	**0.032** **	0.049	**5.465*** **	0.000
		(-4.160)	(0.448)	(0.708)	(1.045)	(2.336)			
Finanztest	1 year	-0.031	0.010	0.004	0.008	0.007	0.003	0.325	0.807
		(-0.993)	(0.267)	(0.116)	(0.199)	(0.153)			
	3 years	**-0.050*** **	0.012	0.017	0.024	**0.040*** **	0.070	**7.905*** **	0.000
		(-4.088)	(0.553)	(0.793)	(1.125)	(2.901)			
FondsNote	1 year	-0.027	0.008	0.001	0.007	0.004	0.002	0.243	0.792
		(-0.753)	(0.200)	(0.028)	(0.170)	(0.089)			
	3 years	**-0.051*** **	0.017	0.018	0.026	**0.040*** **	0.072	**8.167*** **	0.002
		(-6.531)	(0.909)	(0.992)	(1.284)	(4.052)			
Panel B: Geometric mean of monthly returns as alternative predictor and performance measure									
Feri Trust	1 year	**0.124***	-0.002	-0.004	-0.001	-0.003	0.000	0.004	0.457
		(1.854)	(-0.027)	(-0.045)	(-0.013)	(-0.031)			
	3 years	**0.024** **	0.013	0.018	0.025	**0.040** **	0.009	0.944	0.000
		(2.206)	(0.411)	(0.647)	(0.935)	(2.072)			
Finanztest	1 year	**0.130***	-0.006	-0.005	-0.002	0.007	0.000	0.038	0.602
		(1.820)	(-0.068)	(-0.053)	(-0.022)	(0.063)			
	3 years	**0.029*** **	0.010	0.016	0.022	**0.053*** **	0.015	1.650	0.000
		(2.844)	(0.314)	(0.599)	(0.856)	(2.751)			
FondsNote	1 year	**0.134***	-0.008	-0.007	-0.009	-0.004	0.000	0.019	0.446
		(1.897)	(-0.092)	(-0.068)	(-0.088)	(-0.032)			
	3 years	**0.030*** **	0.011	0.016	0.022	**0.050** **	0.013	1.391	0.000
		(3.076)	(0.356)	(0.616)	(0.840)	(2.530)			

Panel C: Carhart four-factor alpha as alternative predictor and geometric mean as performance measure

Feri Trust	1 year	**0.117***	0.004	0.007	0.010	0.005	0.000	0.018	0.218
		(1.665)	(0.043)	(0.068)	(0.093)	(0.050)			
	3 years	**0.030*****	0.010	0.012	0.017	0.026	0.004	0.380	0.000
		(2.598)	(0.341)	(0.433)	(0.645)	(1.227)			
Finanztest	1 year	0.125	0.000	0.002	0.006	0.012	0.000	0.034	0.061
		(1.642)	(-0.002)	(0.022)	(0.059)	(0.105)			
	3 years	**0.037*****	0.006	0.008	0.014	**0.032***	0.006	0.622	0.000
		(3.372)	(0.207)	(0.300)	(0.534)	(1.653)			
FondsNote	1 year	0.126	0.000	-0.001	0.003	0.009	0.000	0.021	0.126
		(1.581)	(0.005)	(-0.005)	(0.029)	(0.081)			
	3 years	**0.040*****	0.005	0.007	0.010	0.028	0.004	0.466	0.000
		(3.482)	(0.184)	(0.240)	(0.379)	(1.358)			

Panel D: Geometric mean as alternative predictor and Carhart four-factor alpha as performance measure

Feri Trust	1 year	-0.026	0.001	-0.003	0.000	-0.002	0.001	0.065	0.657
		(-1.042)	(0.022)	(-0.089)	(0.005)	(-0.047)			
	3 years	**-0.045*****	0.008	0.007	0.012	0.022	0.022	**2.354***	0.221
		(-3.339)	(0.323)	(0.313)	(0.555)	(1.456)			
Finanztest	1 year	-0.029	0.005	0.000	0.006	0.007	0.002	0.231	0.843
		(-1.037)	(0.131)	(-0.001)	(0.162)	(0.156)			
	3 years	**-0.046*****	0.010	0.010	0.016	**0.036****	0.057	**6.343*****	0.082
		(-3.323)	(0.394)	(0.498)	(0.675)	(2.297)			
FondsNote	1 year	-0.028	0.005	0.004	0.009	0.007	0.002	0.203	0.107
		(-0.797)	(0.123)	(0.086)	(0.206)	(0.137)			
	3 years	**-0.049*****	0.014	0.016	0.021	**0.038*****	0.068	**7.681****	0.000
		(-4.669)	(0.614)	(0.903)	(1.121)	(3.032)			

The table presents the results from the monthly regressions of fund performance on the alternative predictor quintiles for the following fund ratings: the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. *t*-statistics are adjusted using the Newey-West procedure with a lag of eleven months or thirty-five months depending on the post-rating period. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively. MR^{All} presents the p-value of the monotonic relation test of Patton and Timmermann (2010).

Table 5.1 points out that the alternative predictors are not able to predict the post-rating performance of funds. Consequently, the future fund performance between quintiles, especially for one-year post-rating periods, does not differ significantly from each other. This is confirmed by the results of the F-statistics that are only significant in 6 out of 24 cases. Interestingly, differences in the post-rating performance between quintiles can only be seen for three-year periods and when the Carhart four-factor alpha is used as performance measure. Nonetheless, the absolute four-factor alpha performance is consistently negative for the formed quintiles and is in line with the performance of the original fund ratings.

Table 5.1 indicates some significant differences in performance between the best and the worst-rated funds if the post-rating performance is measured with the geometric mean. This result can be ascribed to an alpha error accumulation because F-statistics are insignificant. Interestingly, the monotonic relation tests almost reject the null hypothesis if the performance is based

on a three-year post-rating period. Consequently, the coefficients of the dummy variable regression increase as we move from β_2 to β_5.

Further significant differences between quintiles can be seen with the one-factorial ANOVA. For the sake of brevity, this study only presents pairwise comparisons for alternative predictors with significant F-statistics (Table 5.1, Panel A and Panel D). Table 5.2 shows that quintile 5 frequently has a higher three-year post-rating performance than quintile 1, 2 and 3. Instead, the performance of the best and second-best-rated funds does not significantly differ from each other.

Table 5.2: ANOVA – Alternative Predictors

		Feri Trust		Finanztest		FondsNote	
		1 year	3 years	1 year	3 years	1 year	3 years
(I)	(J)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)
Panel A: Carhart four-factor alpha as alternative predictor and performance measure							
β_1	β_2	0.010	0.010	0.010	0.012	0.008	**0.017***
	β_3	0.004	0.015	0.004	**0.017***	0.001	**0.018****
	β_4	0.004	**0.023*****	0.008	**0.024*****	0.007	**0.026*****
	β_5	0.003	**0.032*****	0.007	**0.040*****	0.004	**0.040*****
β_2	β_3	-0.006	0.005	-0.006	0.005	-0.007	0.001
	β_4	-0.006	0.012	-0.002	0.012	-0.001	0.009
	β_5	-0.007	**0.022***	-0.003	**0.028*****	-0.004	**0.023****
β_3	β_4	0.000	0.008	0.003	0.007	0.006	0.007
	β_5	-0.001	0.017	0.002	**0.023****	0.003	**0.022***
β_4	β_5	-0.001	0.009	-0.001	0.016	-0.003	0.015
Panel B: Geometric mean as alternative predictor and four-factor alpha as performance measure							
β_1	β_2	0.001	0.008	0.005	0.010	0.005	0.014
	β_3	-0.003	0.007	0.000	0.010	0.004	0.016
	β_4	0.000	0.012	0.006	0.016	0.009	**0.021****
	β_5	-0.002	**0.022****	0.007	**0.036*****	0.007	**0.039*****
β_2	β_3	-0.004	-0.001	-0.005	0.000	-0.001	0.002
	β_4	0.000	0.004	0.001	0.005	0.004	0.007
	β_5	-0.002	0.013	0.002	**0.026****	0.002	**0.025****
β_3	β_4	0.003	0.005	0.006	0.005	0.005	0.005
	β_5	0.001	0.014	0.007	**0.026****	0.003	**0.023****
β_4	β_5	-0.002	0.010	0.001	**0.020***	-0.002	0.018

The table presents the post-hoc tests of the one-factorial ANOVA on the alternative predictor quintiles based on the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote. Depending on the results of the Levene-statistics, test statistics of post-hoc tests are either adjusted with Tukey (HSD) or Games-Howel. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

Comparing the forecasting ability of the three original fund ratings with the alternative predictors indicates that the post-rating performance of funds with Finanztest-Bewertung can be improved when quintiles are formed with alternative predictors. Consequently, the Finanztest-Bewertung can offer a little value added for investment decisions on German equity funds. By

contrast, Müller and Weber (2014) find a similar forecasting ability between the Finanztest-Bewertung and the alternative predictors. The reason could be that Müller and Weber (2014) only report results for the complete sample, which consists of different equity markets. Moreover, the examination period (2002-2008) differs significantly from that used in this study. Quintiles formed with alternative predictors, which are based on funds with FondsNote, mainly differ from each other in the same manner as with the original fund rating. Only the Feri Trust rating indicates a higher forecasting ability than the alternative predictors used.

Combination of alternative predictors

This section answers the question whether the forecasting ability can be increased when combining the three fund ratings, which are formed with two alternative predictors. As before, quintiles are based on the five-year pre-rating performance measured with the four-factor alpha and the geometric mean of monthly returns. Based on these quintiles, the post-rating performance is analyzed with both performance measures.

Table 6.1 shows that the forecasting ability can be enhanced in some cases when combining the fund ratings of alternative predictors. Fund rating combinations lead to an improvement of predictability if all three fund ratings are combined with each other. Moreover, the combination of funds that are rated with the Feri Trust rating and the Finanztest-Bewertung enhances the forecasting ability of alternative predictors. Nonetheless, analyzed quintiles only significantly differ in performance if the post-rating performance is measured with the four-factor alpha and depends on a three-year period. All other fund rating combinations yield insignificant performance differences between quintiles irrespective of the alternative predictor and the performance measure (see also Table A3 in the Appendix).

Predictability of fund ratings that are formed with alternative predictors is still lower than when combining the original fund ratings. The monotonic relation tests confirm the poor predictability. The null hypothesis of non-increasing regression coefficients cannot be rejected. Moreover, the results are in line with the rating persistence analysis of alternative predictors.

Table 6.1:
Dummy Variable Regression – Combination of Fund Ratings Based on Alternative Predictors

Fund rating		β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.	MR^{All}
Panel A: Carhart four-factor alpha as alternative predictor and performance measure									
All ratings	1 year	-0.041	0.018	0.014	0.017	0.016	0.012	1.245	0.463
		(-1.324)	(0.519)	(0.362)	(0.437)	(0.381)			
	3 years	**-0.061*** **	0.034	**0.045** **	**0.050** **	**0.052*** **	0.082	**9.351*** **	0.000
		(-4.223)	(1.320)	(2.483)	(2.452)	(2.630)			
FTR & FT	1 year	**-0.043***	0.021	0.016	0.018	0.011	0.015	1.597	0.801
		(-1.955)	(0.705)	(0.506)	(0.593)	(0.296)			
	3 years	**-0.072*** **	**0.046***	**0.053*** **	**0.055*** **	**0.048** **	0.095	**11.000*** **	0.794
		(-4.816)	(1.764)	(2.943)	(2.631)	(2.308)			
FTR & FN	1 year	-0.023	0.005	-0.005	0.001	-0.008	0.005	0.560	0.884
		(-0.767)	(0.133)	(-0.125)	(0.026)	(-0.189)			
	3 years	**-0.016** **	-0.003	-0.003	0.002	-0.007	0.004	0.379	0.959
		(-2.477)	(-0.160)	(-0.251)	(0.116)	(-0.572)			
FT & FN	1 year	-0.021	0.003	-0.006	0.000	-0.006	0.003	0.327	0.931
		(-0.593)	(0.086)	(-0.125)	(0.003)	(-0.140)			
	3 years	**-0.014***	-0.003	-0.001	0.001	0.005	0.003	0.266	0.243
		(-1.845)	(-0.190)	(-0.093)	(0.038)	(0.297)			
Panel B: Geometric mean as alternative predictor and Carhart four-factor alpha as performance measure									
All ratings	1 year	**-0.034***	0.006	0.005	0.011	0.012	0.004	0.473	0.037
		(-1.697)	(0.212)	(0.168)	(0.321)	(0.332)			
	3 years	**-0.039*** **	0.000	0.010	0.027	**0.036** **	0.052	**5.784*** **	0.007
		(-3.516)	(-0.012)	(0.464)	(1.443)	(2.036)			
FTR & FT	1 year	**-0.035** **	0.007	0.006	0.010	-0.001	0.006	0.581	0.772
		(-2.279)	(0.300)	(0.224)	(0.353)	(-0.043)			
	3 years	**-0.043*** **	0.001	0.014	0.026	0.023	0.033	**3.567*** **	0.543
		(-4.465)	(0.051)	(0.735)	(1.521)	(1.527)			
FTR & FN	1 year	-0.014	-0.010	-0.012	-0.009	-0.018	0.008	0.834	0.883
		(-0.477)	(-0.295)	(-0.306)	(-0.243)	(-0.407)			
	3 years	**-0.015** **	-0.011	-0.005	-0.004	-0.002	0.005	0.549	0.509
		(-2.055)	(-0.656)	(-0.353)	(-0.269)	(-0.134)			
FT & FN	1 year	-0.018	-0.006	-0.008	-0.004	-0.006	0.002	0.169	0.482
		(-0.540)	(-0.166)	(-0.195)	(-0.100)	(-0.123)			
	3 years	**-0.015** **	-0.008	-0.002	0.002	0.011	0.013	1.366	0.506
		(-2.060)	(-0.447)	(-0.103)	(0.129)	(0.704)			

The table presents the results from the monthly regressions of fund performance on the alternative predictor quintiles for fund rating combinations based on the Feri Trust rating (FTR), the Finanztest-Bewertung (FT), and the FondsNote (FN). Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. *t*-statistics are adjusted using the Newey-West procedure with a lag of eleven months or thirty-five months depending on the post-rating period. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively. MR^{All} presents the p-value of the monotonic relation test of Patton and Timmerman (2010).

Using the four-factor alpha as alternative predictor and performance measure, Table 6.2 does not indicate further performance differences between formed quintiles. However, the dummy variable analysis does not detect all differences in performance. Considering the geometric mean as alternative predictor and the four-factor alpha as three-year performance measure, F-statistics in Table 6.1 indicate further differences between quintiles. Based on the one-factorial ANOVA, the

three-year post-rating performance of the best-rated funds significantly differs from all rating classes when combining all three fund ratings. Only the performance between the best and second-best-rated funds does not differ from each other.

Table 6.2: ANOVA – Combination of Fund Ratings Based on Alternative Predictors

		All ratings		FTR & FT		FTR & FN		FT & FN	
		1 year	3 years	1 year	3 years	1 year	3 years	1 year	3 years
(I)	(J)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)	(J)-(I)
Panel A: Carhart four-factor alpha as alternative predictor and performance measure									
β_1	β_2	0.018	**0.034****	0.021	**0.046*****	0.005	-0.003	0.003	-0.003
	β_3	0.014	**0.045*****	0.016	**0.053*****	-0.005	-0.003	-0.006	-0.001
	β_4	0.017	**0.050*****	0.018	**0.055*****	0.001	0.002	0.000	0.001
	β_5	0.016	**0.052*****	0.011	**0.048*****	-0.008	-0.007	-0.006	0.005
β_2	β_3	-0.004	0.011	-0.005	0.007	-0.010	-0.001	-0.009	0.002
	β_4	-0.001	0.015	-0.003	0.009	-0.004	0.005	-0.003	0.004
	β_5	-0.002	0.018	-0.010	0.002	-0.013	-0.005	-0.010	0.008
β_3	β_4	0.003	0.004	0.002	0.002	0.006	0.005	0.006	0.002
	β_5	0.002	0.007	-0.005	-0.005	-0.003	-0.004	-0.001	0.006
β_4	β_5	-0.001	0.002	-0.007	-0.008	-0.009	-0.009	-0.007	0.004
Panel D: Geometric mean as alternative predictor and Carhart four-factor alpha as performance measure									
β_1	β_2	0.006	0.000	0.007	0.001	-0.010	-0.011	-0.006	-0.008
	β_3	0.005	0.010	0.006	0.014	-0.012	-0.005	-0.008	-0.002
	β_4	0.011	**0.027****	0.010	**0.026****	-0.009	-0.004	-0.004	0.002
	β_5	0.012	**0.036*****	-0.001	**0.023***	-0.018	-0.002	-0.006	0.011
β_2	β_3	0.000	0.011	0.000	0.013	-0.002	0.006	-0.002	0.006
	β_4	0.005	**0.027***	0.004	**0.025***	0.000	0.007	0.002	0.010
	β_5	0.006	**0.036*****	-0.008	0.022	-0.009	0.009	0.001	0.019
β_3	β_4	0.005	0.016	0.004	0.012	0.003	0.001	0.004	0.004
	β_5	0.007	**0.025***	-0.008	0.008	-0.006	0.003	0.002	0.013
β_4	β_5	0.001	0.009	-0.012	-0.004	-0.009	0.002	-0.002	0.009

The table presents the post-hoc tests of the one-factorial ANOVA on the alternative predictor quintiles for fund rating combinations based on the Feri Trust rating (FTR), the Finanztest-Bewertung (FT), and the FondsNote (FN). Depending on the results of the Levene-statistics, test statistics of post-hoc tests are either adjusted with Tukey (HSD) or Games-Howel. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

5 Conclusion

This study identifies the degree to which the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote are able to predict the future fund performance. To my knowledge, no previous research has considered the predictability of three different fund ratings at the same time. Moreover, this study analyzes the predictability of fund rating combinations. The idea behind this approach is that the three examined fund rating agencies evaluate their analyzed funds differently. If the fund ratings assigned are still identical between the rating agencies, the validity of the rat-

ings should be higher than a fund rating that is not confirmed by other fund ratings. The analysis focuses on equity mutual funds investing in Germany.

This study uses two different methods to determine the predictability of examined fund ratings. First, the rating persistence measures the period of time until rated funds change their rating. The longer the period of time is until rating changes can be identified, the higher the predictability of the fund rating agency. Second, this study analyzes the post-rating performance of funds. Significant differences in performance between differently rated funds indicate predictability of fund ratings. Fund performances are measured with the Carhart four-factor alpha and the geometric mean of monthly returns for different post-rating periods.

The analysis leads to the following results. The rating persistence is highest for the FondsNote, followed by the Feri Trust rating and the Finanztest-Bewertung. Nonetheless, funds change their ratings in a short period of time. Interestingly, fund rating combinations do not enhance the period of time until a rated fund changes the rating.

Post-rating differences in performance can only be seen for the Feri Trust rating and the FondsNote. Moreover, fund rating combinations can lead to significant differences in the post-rating performance, but they depend on the combination, the performance measure and the post-rating period. Interestingly, the combination of all three fund ratings yields the highest predictability and contradicts the rating persistence measurements.

Future research may investigate whether these results are stable for other fund ratings and fund samples. In this context, differences in fund rating methodology as described within this study could be analyzed in more detail. Moreover, further influencing factors could be detected that enhance predictability.

References

Antypas, A., Caporale, G. M., Kourogenis, N., Pittas, N. (2009). Selectivity, market timing and the Morningstar star-rating system. DIW Berlin Discussion Paper, No. 874, April.

Blake, C. R., Morey, M. R. (2000). Morningstar ratings and mutual fund performance. The Journal of Financial and Quantitative Analysis 35 (3), pp. 451-483.

Brückner, R., Lehmann, P., Schmidt, M. H., Stehle, R. (2014). Fama/French factors for Germany: which set is best? Working paper, Humboldt University of Berlin, February.

Duret, A., Hereil, P., Mitaine, P., Moussavi, N., Roncalli, T. (2008). Fund rating systems and performance predictability. Working paper, University of Evry, April.

Fama, E. F., MacBeth, J. D. (1973). Risk, return, and equilibrium: empirical tests. The Journal of Political Economy 81 (3), pp. 607-636.

Füss, R., Hille, J., Rindler, P., Schmidt, J., Schmidt, M. (2010). From rising stars and falling angels: on the relationship between the performance and ratings of German mutual funds. Journal of Wealth Management 13 (1), pp. 75-90.

Garnier, O., Pujol, T. (2007). Can today's stars be used to read the stars of the future? Working paper, No. 3, March.

Hereil, P., Mitaine, P., Moussavi, N., Roncalli, T. (2010). Mutual fund ratings & performance persistence. White paper 3, Quant Research by Lyxor, June.

Israel, R. B., Rosenthal, J. S., Wei, J. Z. (2001). Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings. Mathematical Finance 11 (2), pp. 245-265.

Kräussl, R., Sandelowsky, R. (2007). The predictive performance of Morningstar's mutual fund ratings. Working paper, VU University Amsterdam, August.

Meinhardt, C. (2009). Die Aussagekraft von Ratings und Rankings. Die Bank, No. 12, pp. 14-17.

Meinhardt, C. (2014). Fondsbewertungen in Deutschland – Ein Vergleich der Vorhersagekraft. In: dissertation: essays on actively and passively managed financial products, Humboldt University of Berlin, pp. 39-71, May 2015.

Morey, M. R., Gottesman, A. (2006). Morningstar mutual fund ratings redux. Journal of Investment Consulting 8 (1), pp. 25-37.

Müller, S., Weber, M. (2014). Evaluating the rating of Stiftung Warentest: how good are mutual fund ratings and can they be improved? European Financial Management 20 (2), pp. 207-235.

Patton, A. J., Timmermann, A. (2010). Monotonicity in asset returns: new tests with applications to the term structure, the CAPM, and portfolio sorts. Journal of Financial Economics 98 (3), pp. 605-625.

Appendix

Table A1: Monthly Transition Matrices

Feri Trust Rating

	Q5	Q4	Q3	Q2	Q1	NR
Q5	0.9106	0.0878	0	0	0	0.0015
Q4	0.0328	0.8945	0.0691	0	0	0.0036
Q3	0	0.0530	0.8732	0.0665	0	0.0073
Q2	0	0	0.0647	0.8882	0.0344	0.0127
Q1	0	0	0	0.1284	0.8581	0.0135
NR	0.0028	0.0014	0.0057	0.0043	0.0021	0.9837

Finanztest

	Q5	Q4	Q3	Q2	Q1	NR
Q5	0.8735	0.1206	0	0	0	0.0059
Q4	0.0273	0.8401	0.1248	0.0007	0	0.0072
Q3	0.0004	0.0755	0.8384	0.0722	0.0004	0.0131
Q2	0	0.0005	0.1041	0.8467	0.0361	0.0126
Q1	0	0	0	0.1394	0.8446	0.0159
NR	0.0003	0.0052	0.0069	0.0066	0.0024	0.9785

FondsNote

	Q5	Q4	Q3	Q2	Q1	NR
Q5	0.9394	0.0572	0	0	0	0.0034
Q4	0.0165	0.9414	0.0384	0	0	0.0037
Q3	0	0.0332	0.9368	0.0286	0	0.0014
Q2	0	0.0006	0.0416	0.9409	0.0149	0.0019
Q1	0	0	0	0.0538	0.9405	0.0057
NR	0.0059	0.0125	0.0066	0.0081	0.0029	0.9640

The table presents the results of the empirical monthly transition matrix based on the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote (examination period: 12/2002-12/2009). The best-rated to the worst-rated funds are represented by quintile 5 (Q5) to 1 (Q1). The class "NR" states unrated funds that are included or excluded from the universe of rated funds due to liquidation, mergers or changes in investment focus.

Table A2.1: Fund Rating Persistence – Original Fund Ratings

	Original fund rating			Fund rating combination			
	Feri Trust	Finanztest	FondsNote	All ratings	Feri Trust & Finanztest	Feri Trust & FondsNote	Finanztest & FondsNote
Q1	4.96	4.46	12.40	5.63	5.84	7.19	5.66
Q2	8.42	5.40	15.16	2.35	2.87	3.92	3.47
Q3	6.54	5.63	14.89	2.84	2.97	5.20	3.72
Q4	8.34	5.05	16.39	2.79	3.12	5.66	3.68
Q5	8.59	5.54	12.35	4.23	4.49	12.05	3.29

The table presents the period of time (in months) until rated funds change their rating with a probability of 50%. The period of time is separately measured for different rating classes. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0. Fund rating combinations are based on the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote.

Table A2.2: Fund Rating Persistence – Alternative Predictors

	Carhart four-factor alpha			Geometric mean of monthly returns		
	Feri Trust	Finanztest	FondsNote	Feri Trust	Finanztest	FondsNote
Q1	6.93	6.83	8.28	7.71	7.39	8.30
Q2	2.88	3.10	3.10	2.64	2.79	3.31
Q3	2.75	2.75	2.76	2.51	2.69	2.97
Q4	3.60	3.52	3.23	3.64	3.90	3.53
Q5	10.25	9.55	9.28	10.46	11.33	10.41

The table presents the period of time (in months) until rated funds change their rating with a probability of 50%. The period of time is separately measured for different rating classes. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0.

Table A2.3: Fund Rating Persistence – Fund Rating Combinations with Alternative Predictors

	Carhart four-factor alpha				Geometric mean of monthly returns			
	All ratings	Feri Trust & Finanztest	Feri Trust & FondsNote	Finanztest & FondsNote	All ratings	Feri Trust & Finanztest	Feri Trust & FondsNote	Finanztest & FondsNote
Q1	6.73	6.26	7.05	8.58	5.93	6.32	6.12	8.54
Q2	2.24	2.66	2.22	2.70	1.79	2.24	1.79	3.10
Q3	1.69	2.08	1.72	2.09	1.31	1.79	1.45	1.91
Q4	1.93	2.47	2.01	2.12	1.70	2.13	1.85	2.11
Q5	5.13	5.72	5.04	5.82	5.56	6.53	5.53	5.63

The table presents the period of time (in months) until rated funds change their rating with a probability of 50%. The period of time is separately measured for different rating classes. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0. Fund rating combinations are based on the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote.

Table A3: Dummy Variable Regression – Combination of Fund Ratings Based on Alternative Predictors

Fund rating		β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.	MR^{All}
Panel C: Geometric mean of monthly returns as alternative predictor and performance measure									
All ratings	1 year	**0.129***	-0.006	-0.005	-0.002	0.004	0.000	0.021	0.515
		(1.866)	(-0.068)	(-0.052)	(-0.022)	(0.034)			
	3 years	**0.076*****	0.001	0.006	0.006	0.014	0.002	0.164	0.031
		(5.548)	(0.022)	(0.218)	(0.178)	(0.559)			
FTR & FT	1 year	**0.123***	-0.001	-0.002	0.004	-0.007	0.000	0.020	0.785
		(1.813)	(-0.011)	(-0.019)	(0.035)	(-0.065)			
	3 years	**0.073*****	0.003	0.007	0.007	0.002	0.001	0.055	0.562
		(5.542)	(0.077)	(0.252)	(0.230)	(0.081)			
FTR & FN	1 year	**0.126***	-0.001	0.001	0.005	-0.009	0.000	0.037	0.851
		(1.800)	(-0.012)	(0.015)	(0.049)	(-0.085)			
	3 years	**0.073*****	0.003	0.010	0.008	0.004	0.001	0.075	0.351
		(5.470)	(0.084)	(0.356)	(0.238)	(0.165)			
FT & FN	1 year	**0.133***	-0.006	-0.004	-0.003	0.001	0.000	0.011	0.301
		(1.834)	(-0.070)	(-0.035)	(-0.031)	(0.010)			
	3 years	**0.078*****	0.002	0.007	0.005	0.013	0.001	0.135	0.129
		(5.543)	(0.052)	(0.258)	(0.167)	(0.525)			
Panel D: Carhart four-factor alpha as alternative predictor and geometric mean as performance measure									
All ratings	1 year	0.120	0.002	0.006	0.013	0.014	0.001	0.052	0.051
		(1.592)	(0.022)	(0.055)	(0.120)	(0.125)			
	3 years	**0.073*****	-0.002	0.010	0.012	0.015	0.003	0.300	0.201
		(5.011)	(-0.054)	(0.353)	(0.376)	(0.619)			
FTR & FT	1 year	**0.117***	0.004	0.006	0.010	0.010	0.000	0.027	0.002
		(1.664)	(0.039)	(0.063)	(0.101)	(0.095)			
	3 years	**0.070*****	0.002	0.007	0.009	0.005	0.001	0.079	0.135
		(5.170)	(0.042)	(0.254)	(0.274)	(0.215)			
FTR & FN	1 year	0.121	0.002	0.003	0.011	0.007	0.000	0.023	0.060
		(1.632)	(0.027)	(0.034)	(0.102)	(0.065)			
	3 years	**0.075*****	-0.002	0.005	0.006	0.001	0.001	0.064	0.196
		(5.413)	(-0.060)	(0.180)	(0.183)	(0.040)			
FT & FN	1 year	0.128	-0.002	-0.002	0.005	0.012	0.000	0.046	0.173
		(1.561)	(-0.021)	(-0.018)	(0.045)	(0.103)			
	3 years	**0.077*****	-0.003	0.005	0.005	0.012	0.002	0.165	0.204
		(5.146)	(-0.074)	(0.172)	(0.147)	(0.466)			

The table presents the results from the monthly regressions of fund performance on the alternative predictor quintiles for fund rating combinations based on the Feri Trust rating (FTR), the Finanztest-Bewertung (FT), and the FondsNote (FN). Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. *t*-statistics are adjusted using the Newey-West procedure with a lag of eleven months or thirty-five months depending on the post-rating period. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively. MR^{All} presents the p-value of the monotonic relation test of Patton and Timmerman (2010).

Part V:
Evaluating the Mutual Fund Rating of Axel Springer and FondsConsult: A Case Study

Abstract

This study addresses the question of whether qualitative factors can improve the predictability of fund ratings to distinguish between well and poorly performing funds in the future. As qualitative factors are similar between fund rating agencies, we can examine the value added of one fund rating using these factors. Analyses are based on the FondsNote, a German fund rating jointly provided by Axel Springer and FondsConsult. Besides qualitative factors, the fund's costs and the behavior of investors on fund ratings are analyzed as further influencing factors.

It is shown that the overall predictability of the FondsNote is low, but it depends on the fund rating category analyzed and the performance metric. Qualitative factors hardly improve the predictability of the FondsNote. However, the costs and behavior of investors on fund ratings significantly influence the ability to predict the future performance.

JEL Classification: G1, G11, G12, G14

Keywords: predictability, qualitative and quantitative factors, rating and performance persistence

1 Introduction

Mutual funds are rated by several fund rating agencies. Each of them aims to reduce extensive information to one quality label. Doing so helps the investors in their decision-making. The most well-known mutual fund rating agency worldwide is Morningstar with its five-star rating system. Standard & Poor's, Feri Trust, Fitch, Lipper, Sauren, and Scope are some other fund rating agencies that evaluate mutual funds.

Literature shows that investors combine these mutual fund ratings with high performance expectations (e.g. Sirri and Tufano (1998), Del Guercio and Tkac (2008), Füss et al. (2010), and Meinhardt (2014a)). In this context, the question arises whether investors should include fund ratings in their decision-making. This obviously depends on whether the mutual fund ratings are indeed able to predict the future performance of the funds examined.

It appears that fund rating agencies that only include backward-looking quantitative factors in their evaluation process can barely meet the investors' requirements (e.g. Blake and Morey (2000), Duret et al. (2008), Antypas et al. (2009), Füss et al. (2010), and Meinhardt (2014a)). A solution could be to consider the qualitative factors, which suggest predictive power (e.g. BVI (2007a) and Meinhardt (2014a)). This study deals with this topic by analyzing the FondsNote, a mutual fund rating provided by Axel Springer and FondsConsult, which comprises both quantitative and qualitative factors. This way, we can analyze in detail the effective influence of qualitative factors on the final mutual fund rating.

Furthermore, if qualitative factors are included in the evaluation process, they are often similar between fund rating agencies. Only the weight of qualitative factors within the final fund rating differs significantly from each other. For instance, Axel Springer and FondsConsult analyze the continuity of the fund management, consistency of the investment approach, and truth and clarity of the investment product (Axel Springer and FondsConsult (2007)). Similar qualitative factors are used by Feri Trust, Standard & Poor's, Sauren, and Scope (BVI (2007b)). Consequently, the findings in this study for the FondsNote can be applied to other mutual fund ratings as well.

Besides the analyses of the influence of qualitative factors on the predictability of fund ratings, this study also focuses on other influencing factors. These include the costs of mutual funds and the behavior of investors on fund ratings. If the costs significantly differ between funds and fund rating classes, they influence the funds' returns differentially (e.g. Elton et al. (1993), Carhart (1997), and Baker et al. (2009)) and, therefore, the predictability of the fund rating. The same is true for the behavior of investors on fund ratings. For instance, if the best-rated funds are faced with disproportionately large inflows compared to other fund rating classes, these funds can find

it difficult to maintain their past performance (e.g. Chen et al. (2004)). Consequently, the predictability of the fund rating will also go down in the future.

This study is divided into four sections. Following a brief introduction, Section 2 presents the dataset and the methodology used in this study. Section 3 discusses the results and Section 4 summarizes the findings.

2 Data and Methodology

2.1 Data

Axel Springer and FondsConsult publish their fund rating "FondsNote" on a monthly basis. The FondsNote is available for the period from December 2002 to December 2009. We receive data containing the following variables: fund name, WKN, ISIN, fund category, FondsNote, qualitative sub-rating, and reporting date. These data comprise all the mutual funds rated within the sample period. Therefore, the dataset is free of the problem of survivorship-bias because it includes both dead (liquidated or merged) as well as surviving funds.

Fund return data is computed by using the total return index from Thomson Reuters Datastream. The total return index (code: RI) measures the net asset value of any fund by assuming reinvested dividends (net of fees). This way, the funds that distribute dividends can easily be compared with accumulating funds. Fund cost data is also obtained from Thomson Reuters Datastream (code: TER) by using the total expense ratio (TER). These data are cross-checked and complemented with weekly TERs published by Fondsxpress to eliminate missing values and to offset potential mistakes as well. TER measures the ratio of the total cost of the fund and the fund's total assets on an annual basis. A fund's total cost is inclusive of management fees, but does not comprise transactional costs, performance fees and sales loads. Fund total assets are gathered from Bloomberg because these data are not available from Thomson Reuters Datastream. Following earlier studies on mutual fund ratings (e.g. Blake and Morey (2000), Del Guercio and Thak (2008), and Müller and Weber (2014)), this study also includes all fund share classes. The results show that fund ratings generally vary among the fund share classes of the same fund. Fund returns, fund costs as well as fund total assets are available on a monthly basis across the entire sample period.

I analyze the overall predictability of the FondsNote, as well as for seven different fund rating categories. These include the following five equity fund categories: World, North America, Europe, Germany, and Japan. Moreover, two fixed-income fund categories are examined: World and EUR. These are the largest equity and fixed-income fund categories. Thus we can assume that

these fund categories receive the maximum attention from investors. Funds like fund of funds, real estate funds and multi asset funds are not analyzed separately because of the low number of rated funds within these categories. No meaningful results can be obtained by analyzing these small categories. Nonetheless, these funds are included in the overall examination of the predictability of the FondsNote.

Table 1 illustrates the total number of rated funds for the different fund categories in December of each year from 2002 to 2009. Within the sample period, 4,848 different funds receive a FondsNote rating. The number of rated funds sharply rises over time, which is in line with the total growth in the industry. Only the fund universe of the equity category "Germany" has remained stable over time.

Table 1: Descriptive Statistics

			Equity					Fixed income	
	Categories	All funds	World	North America	Europe	Germany	Japan	World	EUR
12/2002	29	1,365	123	95	168	95	64	91	131
12/2003	38	1,911	180	140	227	98	82	118	163
12/2004	49	2,645	239	169	272	109	100	136	191
12/2005	53	2,995	270	204	298	107	113	143	195
12/2006	53	3,019	285	208	303	104	111	133	198
12/2007	53	3,081	299	212	317	99	112	132	197
12/2008	62	3,469	305	242	351	105	129	118	214
12/2009	63	3,402	296	231	347	100	127	106	216
In all	65	4,848	424	336	464	151	173	185	286

This table presents both the number of funds rated by Axel Springer and FondsConsult and the total number of fund rating categories in December of each year between 2002 and 2009. The table shows the total number of rated funds as well as the number of funds for seven different fund rating categories. These include the following equity and fixed-income (FI) fund rating categories: World, North America, Europe, Germany, Japan, World (FI), and EUR (FI). Moreover, the total number of funds rated and the total number of fund rating categories is also presented for the entire sample period from December 2002 to December 2009.

Compared to Meinhardt (2014a) and Meinhardt (2014b), who also analyzes the predictability of the FondsNote, differences in the number of rated funds of the equity category "Germany" can be explained by the broader fund universe used in this study. Here, the fund rating category also includes mid-cap funds as well as small-cap funds. The number of fund rating categories also sharply rises over time. At the end of 2002, Axel Springer and FondsConsult had sorted funds into 29 different fund categories as compared to the jump to 63 fund categories at the end of 2009. This study includes 65 different fund rating categories in total.

2.2 Methodology

Dummy variable regression and Markov chain analysis

Literature principally uses two different methodologies to analyze the predictability of fund ratings. First, fund rating predictability is defined as the stability of a fund rating over time (e.g. Garnier and Pujol (2007), Duret et al. (2008), Hereil et al. (2010), and Meinhardt (2014b)). If a fund rating does not change in the short term, then the fund rating assigned in the past is also valid for the present. Consequently, the fund rating possesses predictability.

Literature analyzes fund rating stability by estimating empirical transition matrices. First, funds are sorted into quintiles based on their fund rating in month *t* over the entire sample period. Non-evaluated funds, which will be rated in the future or were rated in the past, are frequently considered as a sixth fund rating class. Using the maximum likelihood method, monthly probabilities are determined which measure whether funds change their fund rating or stay stable. In the Appendix, Table A1 presents the empirical transition matrix for the FondsNote.

In the context of a Markov chain setting, transition matrices are estimated for different periods of time using a valid Markov generator (Israel et al. (2001)). In accordance with Meinhardt (2014b), the stability of a fund rating is measured as the period of time until the rated funds change their rating with a probability of 50%. This way, the results are easily comparable to those in the study of Meinhardt (2014b).

Second, fund rating predictability can be defined as the ability to identify funds that will perform well and poorly in the future. Therefore, funds are sorted into quintiles based on their fund rating each month within the sample period. Then, the post-rating performance is measured by using different performance metrics. If the post-rating performance significantly differs between the fund rating classes, the examined fund rating can predict the future performance of the funds.

Literature uses a dummy variable regression, first introduced by Blake and Morey (2000), to measure performance differences of fund rating classes compared to a reference class (see also Gerrans (2006), Kräussl and Sandelowsky (2007), Antypas et al. (2009), Füss et al. (2010), Müller and Weber (2014), Meinhardt (2014a), and Meinhardt (2014b)). This reference class is generally formed from the best-rated funds or the worst-rated funds. Consequently, significant differences in performance between fund rating classes should be significantly positive or negative depending on the definition of the reference class. This study uses the worst-rated funds as the reference class.

Moreover, this study extends the methodology of Blake and Morey (2000). In line with Müller and Weber (2014) and Meinhardt (2014b), this study analyzes the relationship between fund rating classes and the post-rating performance via multiple dummy variable regressions, using the

procedure of Fama and MacBeth (1973). This way, we can investigate the overall predictability of the fund rating.

Performance measures

Several performance metrics are used in the literature to evaluate the out-of-sample performance of rated funds. For example, Blake and Morey (2000) and Morey and Gottesman (2006) use both the Sharpe ratio and the Jensen alpha as performance metrics. Furthermore, Füss et al. (2010) and Meinhardt (2014a) apply the geometric mean of monthly returns. Müller and Weber (2014) calculate the performance of funds with the index-adjusted return. However, the research results are hardly comparable because of the frequent use of different performance metrics.

In this study, the performance of the post-rating period is measured with all the four performance metrics. We use the geometric mean of monthly returns (*GM*), the Sharpe ratio (*SR*), the index-adjusted return (*IaR*), and the Jensen alpha (α). Thus the performance of fund *i* is calculated in the following way:

$$GM_i = \prod_{t=1}^{T} \left(1 + r_{i,t}\right), \tag{1}$$

$$SR_i = \frac{\overline{r_{i,t} - r_{f,t}}}{\sigma_i}, \tag{2}$$

$$IaR_i = \frac{i}{T}\left(\sum_{t=1}^{T} \left(r_{i,t} - r_{index,t}\right)\right), \tag{3}$$

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i \cdot \left(r_{index,t} - r_{f,t}\right) + \varepsilon_{i,t}. \tag{4}$$

In equation (1) to (4), $r_{i,t}$, $r_{f,t}$, and $r_{index,t}$ correspond to the returns of fund *i*, the risk-free rate, and the index of fund *i* at time *t*. The return of the risk-free rate is the one-month EURIBOR. Index returns are calculated by using Morgan Stanley Capital International (MSCI) index prices for equity funds and Barclays index prices for fixed-income funds. The selected corresponding indices for the different fund rating categories are: the MSCI World, MSCI North America, MSCI Europe, MSCI Germany, MSCI Japan, Barclays Global Aggregate Index, and Barclays Euro Aggregate Index. Index returns are also calculated using STOXX indices and iBoxx indices. For the sake of brevity, the results based on these indices are not shown because they barely differ from the results presented. Thus the analyses are robust for using different indices.

Although the evaluation process of the FondsNote is not only dependent on past performance, these four performance metrics are used as alternative predictors for future performance. This way, we can analyze whether FondsNote's qualitative factors can increase the ability to predict the future performance of funds. Moreover, we can study the stability of fund returns over time. This can be seen as a performance persistence measurement. The pre-rating period comprises 48 months because Axel Springer and FondsConsult use the same time period to calculate their qualitative factors.

Qualitative factors, costs, and the behavior of investors

Fund rating predictability is influenced by various factors. This study examines three different factors: qualitative factors as part of the evaluation process of the FondsNote, the costs of funds, and the changes of funds' inflows and outflows depending on current fund ratings.

First, if qualitative factors are included in the evaluation process of funds, the predictability of the fund rating should increase compared to a fund rating that only relies on quantitative factors. The reason is that quantitative factors like the fund's performance are based only on historical data (BVI (2007a)). However, qualitative factors, for instance, evaluate the current fund management and the fund's investment process. According to the extant literature, frequent fund management turnovers can significantly change the performance of funds (e.g. Bessler et al (2010), Barelkowsa (2010) and Huij et al. (2012)). Management turnovers of winning funds decrease the future performance of the particular funds whereas losing funds benefit from such turnovers. Moreover, Brown et al. (2009) show that a consistent style plays an important role for persistence in performance. Over time, style-consistent funds outperform less style-consistent funds.

The evaluation process of Axel Springer and FondsConsult includes quantitative factors for each of the funds that were analyzed. Funds are also qualitatively analyzed if they receive at least the third-best quantitative sub-rating. Consequently, the FondsNote only includes a quantitative sub-rating or consists of both a quantitative and a qualitative sub-rating as well. A detailed description of the methodology can be found in the study of Meinhardt (2014a). Both the FondsNote and the qualitative sub-rating are available for each fund. Consequently, the quantitative sub-rating can be derived from the other two fund ratings. In this manner, this study is able to analyze the influence of the qualitative sub-rating on the quantitative sub-rating as well as the FondsNote.

Second, Meinhardt (2014a) shows that the costs of funds influence the predictability of the Feri Trust rating, the Finanztest-Bewertung, and the FondsNote by using the fund rating category "Germany" as the underlying fund sample. Similar results are suggested by Müller and Weber

(2014) for the Finanztest-Bewertung by analyzing six different equity fund rating categories. Do these results also apply for the overall FondsNote and their fund rating categories?

Costs are defined as the annualized TER of funds. If TERs do not differ from each other, one can assume that the costs reduce the returns of funds in a similar manner within fund rating classes and, therefore, do not influence the predictability of fund ratings. In contrast, TERs affect funds returns as well as predictability if the costs are not similar across the fund rating classes. To investigate this issue, the funds are ranked on the basis of the FondsNote, their sub-ratings and the four-year pre-rating performance measured with the four performance metrics and then sorted into quintiles, respectively. This way, the monthly total expense ratios can be determined for each quintile.

Third, the relationship between fund ratings and investment behavior can influence the predictability of fund ratings. Literature shows that investors distinguish between well and poorly performing funds on the basis of fund ratings (e.g. Sirri and Tufano (1998), Del Guercio and Tkac (2008), and Füss et al. (2010)). Top-rated funds receive disproportionately large inflows from investors in contrast to the worst-rated funds, which are faced with high outflows.

The challenge of funds, which frequently have high monthly inflows, is to continuously find new investments with high yields. Otherwise, the performance of funds would decrease because inflows cannot go on being infinitely invested into only one asset. On the other hand, funds with high monthly outflows must continuously hold back cash so that investors can sell their fund shares any time. This reduces the funds' overall performance because the accrued income cannot be invested in assets with higher yields. Consequently, if fund ratings indeed change the amount of monthly fund flows, this will also influence the predictability of the fund ratings.

The fund flow is defined as the difference of fund class assets between two consecutive months and is measured for each rated fund within the sample period (Sirri and Tufano (1998)). Just like before, the funds are then sorted into quintiles based on the their fund rating, which include the FondsNote, the sub-ratings of the FondsNote and the alternative predictors, to measure the arithmetic mean of monthly fund flows of each fund rating class.

3 Results

3.1 Fund Rating Predictability

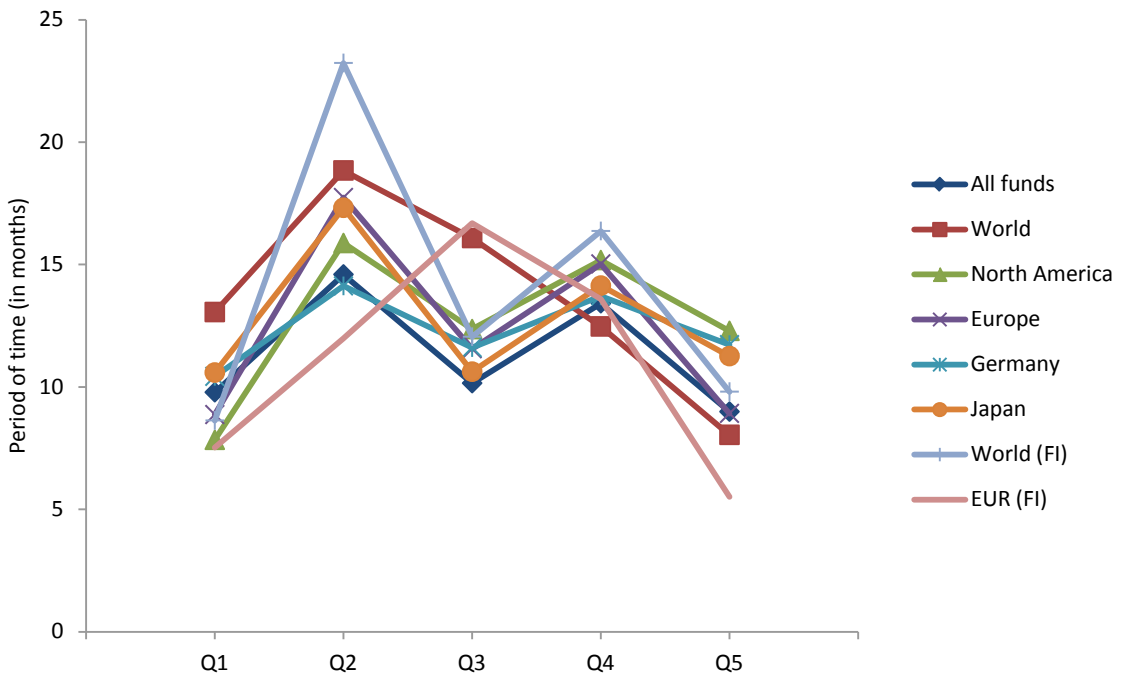
Markov chain analysis

Predictability of the FondsNote can be analyzed in accordance with the stability shown by a given fund rating over time. The longer the period of time when the fund rating remains un-

changed, the higher is the value assigned to the FondsNote predictability. Figure 1 and Table 2 illustrate the results of rating persistence in the context of all the examined fund rating categories.

FondsNote fund ratings persist over a long period of time. Moreover, the fund rating's stability differs among the fund rating classes and categories as well. The best-rated and the worst-rated funds (Q5 and Q1) frequently have the lowest stability. The stability of fund rating goes up when funds get the second-best and the second-worst fund ratings (Q4 and Q2). Moreover, the difference in stability between fund rating classes and fund rating categories can be very high. For example, the best-rated fixed-income funds of the category "EUR" already change their fund ratings after 5.51 months. In contrast, the stability of the second-worst-rated funds of the fixed-income category "World" is four-times higher and amounts to 23.23 months.

Figure 1: Rating Persistence



This figure presents the period of time (left axis: in months) until rated funds change their rating with a probability of 50%. The period of time is separately measured for different rating classes. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0. The category "All funds" includes 65 fund rating categories evaluated by the FondsNote within the sample period from December 2002 to December 2009. The equity fund rating categories examined are: World, North America, Europe, Germany, and Japan. Fixed-income categories (FI) are: World and EUR.

The results are in line with the study of Meinhardt (2014b). In this study, fund ratings persist between 12.35 months and 16.39 months for the same sample period. Compared to the Feri Trust rating and the Stiftung Warentest rating analyzed by Meinhardt (2014b), the stability of the FondsNote persists for a much longer time. Duret et al. (2008) focus on the rating persistence of

the Morningstar rating from 2000 to 2006. The stability of the Morningstar rating is much lower than that of the FondsNote. The results show that the best-rated funds already change their ratings after 4.33 months. Meanwhile, the worst-rated funds only stay stable for 5.74 months. All the other fund rating classes yield similar results. Hereil et al. (2010) analyze the predictability of the Morningstar rating for the period from 2000 to 2009, and use the same rating persistence measure. This study measures the stability of the fund rating for different equity and fixed-income fund rating categories. Funds of the categories "Europe", "US", and "Emerging Markets" quickly change their fund ratings within six months. The same is true for the fixed-income categories. Fund ratings of the categories "US" and "Emerging Markets" do not even stay stable for three months. Consequently, there is some support to the fact that that the qualitative factors can increase the predictability of the FondsNote as compared to the only backward-looking Morningstar rating.

Table 2: Rating Persistence

	All funds	World	North America	Europe	Germany	Japan	World (FI)	EUR (FI)
Q1	9.78	13.06	7.84	8.87	10.43	10.58	8.63	7.52
Q2	14.59	18.84	15.88	17.74	14.12	17.32	23.23	11.99
Q3	10.15	16.08	12.36	11.56	11.60	10.62	12.05	16.69
Q4	13.42	12.47	15.18	15.02	13.72	14.14	16.37	13.60
Q5	8.99	8.04	12.29	8.92	11.72	11.26	9.80	5.51

This table presents the period of time (in months) until rated funds change their rating with a probability of 50%. The period of time is separately measured for different rating classes. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0. The category "All funds" includes all 65 fund rating categories valuated by the FondsNote within the sample period from December 2002 to December 2009. The equity fund rating categories examined are: World, North America, Europe, Germany, and Japan. Fixed-income categories (FI) are: World and EUR.

Dummy variable regression

Tables 3 and 4 present the results of the dummy variable regressions for one-year and three-year post-rating periods. This study uses the geometric mean of monthly returns, the Sharpe ratio, the index-adjusted return, and the Jensen alpha as a benchmark. The results for the first two performance metrics are given in Tables 3 and 4. The results related to the latter two performance metrics are presented in Tables A2 and A3 in the Appendix.

The results show that predictability significantly depends on the performance metric and the post-rating period. Moreover, predictability differs between the fund rating categories. Predictability is low when using the geometric mean of monthly returns and the Sharpe ratio as performance metrics, irrespective of the period of time analyzed. Only the best-rated funds significantly differ in performance from the worst-rated funds using the Sharpe ratio as performance metric and a post-rating period of three years. This can be seen in Table 4 and it holds true for three out

of eight cases. There is also some support for the fact that the FondsNote can separate between well and poorly performing fixed-income funds, though it is only for the fund rating category "World". All the other fund rating categories yield insignificant results.

Table 3:
Performance Predictability of the FondsNote by Using the Geometric Mean and the Sharpe Ratio as Performance Metrics over the next 12 Months

Fund sample	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
Panel A: Geometric mean of monthly returns over next 12 months							
All funds	0.0712	-0.0053	-0.0084	-0.0093	-0.0034	0.00	0.048
	(1.343)	(-0.085)	(-0.124)	(-0.133)	(-0.050)		
World	0.0255	0.0290	0.0250	0.0385	0.0437	0.01	0.708
	(0.462)	(0.446)	(0.328)	(0.492)	(0.473)		
North America	0.0266	0.0001	0.0013	0.0067	0.0179	0.00	0.158
	(0.573)	(0.001)	(0.020)	(0.094)	(0.245)		
Europe	0.0803	-0.0079	-0.0091	0.0035	-0.0009	0.00	0.055
	(1.585)	(-0.104)	(-0.106)	(0.037)	(-0.010)		
Germany	0.0161	0.0036	0.0015	0.0045	0.0236	0.00	0.462
	(0.276)	(0.053)	(0.021)	(0.059)	(0.340)		
Japan	**0.1246***	-0.0147	-0.0169	-0.0154	-0.0087	0.00	0.073
	(1.930)	(-0.183)	(-0.179)	(-0.154)	(-0.084)		
World (FI)	**0.0364*****	-0.0126	-0.0093	-0.0093	-0.0131	0.02	**2.283***
	(3.009)	(-0.803)	(-0.629)	(-0.666)	(-0.919)		
EUR (FI)	0.0311	-0.0086	-0.0061	-0.0063	-0.0079	0.00	0.294
	(1.086)	(-0.237)	(-0.173)	(-0.198)	(-0.260)		
Panel B: Sharpe ratio over next 12 months							
All funds	**0.1801***	0.0511	0.0996	0.1074	0.1422	0.02	**2.583****
	(1.657)	(0.375)	(0.672)	(0.706)	(1.082)		
World	0.0849	0.0457	0.0948	0.1223	0.1558	0.03	**2.797****
	(0.770)	(0.331)	(0.623)	(0.794)	(0.930)		
North America	0.0585	0.0167	0.0319	0.0517	0.0682	0.01	0.725
	(0.807)	(0.163)	(0.283)	(0.416)	(0.594)		
Europe	**0.1736****	0.0427	0.0610	0.0912	0.1363	0.01	1.474
	(1.969)	(0.283)	(0.361)	(0.496)	(0.769)		
Germany	-0.0183	0.0098	0.0229	0.0205	0.0633	0.01	0.997
	(-0.208)	(0.086)	(0.199)	(0.162)	(0.571)		
Japan	**0.2146****	0.0310	0.0338	0.0365	0.0695	0.00	0.411
	(2.252)	(0.236)	(0.235)	(0.246)	(0.438)		
World (FI)	**0.2719*****	0.0028	0.0684	0.0947	0.1082	0.03	**2.713****
	(3.153)	(0.028)	(0.601)	(0.926)	(1.005)		
EUR (FI)	0.0584	0.0006	0.0435	0.1436	**0.2703*****	0.13	**16.051*****
	(0.682)	(0.005)	(0.385)	(1.289)	(2.798)		

This table illustrates the results from the monthly dummy variable regressions of fund performance on the rating quintiles by using the following performance measures: the geometric mean of monthly returns and the Sharpe ratio. The performance is measured over the next 12 months after a given fund rating. Funds are rated monthly from December 2002 to December 2009. The post-rating performance is measured between December 2002 and December 2010. The category "All funds" includes all the 65 FondsNote fund rating categories. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. The *t*-statistics are adjusted using the Newey-West procedure with a lag of 11 months. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

Predictability increases when using the index-adjusted return as performance metric. In particular, rating classes often differ in performance when using a three-year post-rating period.

Nonetheless, better rated funds frequently have a lower performance than the worst-rated funds. For example, Table A3 shows that the worst-rated funds in the category "World" have the highest performance. However, the best-rated funds frequently have the highest three-year post-rating performance.

Table 4:
Performance Predictability of the FondsNote by Using the Geometric Mean and the Sharpe Ratio as Performance Metrics over the next 36 Months

Fund sample	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
Panel A: Geometric mean of monthly returns over next 36 months							
All funds	**0.0482*** **	-0.0112	-0.0104	-0.0106	-0.0045	0.00	0.315
	(5.743)	(-0.587)	(-0.698)	(-0.662)	(-0.297)		
World	**0.0366*** **	-0.0165	-0.0166	-0.0076	0.0009	0.01	0.767
	(2.923)	(-0.779)	(-0.960)	(-0.408)	(0.035)		
North America	**0.0230* **	-0.0121	-0.0117	-0.0094	0.0106	0.01	1.147
	(1.788)	(-0.645)	(-0.614)	(-0.494)	(0.724)		
Europe	**0.0380** **	-0.0062	-0.0110	-0.0039	0.0035	0.00	0.263
	(2.511)	(-0.165)	(-0.451)	(-0.119)	(0.142)		
Germany	**-0.0104*** **	0.0099	0.0067	0.0046	0.0088	0.01	0.531
	(-2.836)	(0.812)	(0.740)	(0.502)	(0.586)		
Japan	**0.0574*** **	-0.0148	0.0040	0.0058	0.0125	0.01	0.744
	(6.724)	(-0.579)	(0.173)	(0.201)	(0.551)		
World (FI)	**0.0123*** **	-0.0001	**0.0071** **	0.0041	0.0048	0.05	**5.833*** **
	(4.948)	(-0.014)	(2.008)	(1.009)	(1.246)		
EUR (FI)	**0.0235** **	-0.0062	-0.0028	-0.0051	-0.0084	0.01	1.084
	(2.083)	(-0.383)	(-0.213)	(-0.368)	(-0.695)		
Panel B: Sharpe ratio over next 36 months							
All funds	**0.1248*** **	0.0188	0.0450	0.0492	0.0585	0.02	1.637
	(8.257)	(0.431)	(1.153)	(1.032)	(1.534)		
World	**0.0583*** **	0.0122	0.0305	0.0520	0.0754	0.02	**2.487** **
	(2.974)	(0.298)	(0.819)	(1.324)	(1.360)		
North America	**0.0515*** **	-0.0025	0.0044	0.0209	**0.0498** **	0.02	1.834
	(3.163)	(-0.084)	(0.147)	(0.709)	(2.447)		
Europe	**0.0930*** **	0.0187	0.0153	0.0417	0.0721	0.01	1.507
	(4.310)	(0.212)	(0.264)	(0.520)	(1.055)		
Germany	**-0.0205*** **	0.0192	0.0146	0.0122	0.0185	0.01	0.689
	(-2.916)	(1.110)	(0.918)	(0.921)	(0.742)		
Japan	**0.1157*** **	-0.0085	0.0289	0.0401	0.0618	0.02	**2.081* **
	(7.785)	(-0.173)	(0.559)	(0.598)	(1.260)		
World (FI)	**0.1001*** **	0.0293	**0.0841*** **	**0.0670*** **	**0.0990*** **	0.14	**17.021*** **
	(6.319)	(1.137)	(4.801)	(2.690)	(3.656)		
EUR (FI)	**0.0657** **	-0.0069	0.0235	0.0605	**0.1191*** **	0.18	**23.473*** **
	(2.458)	(-0.151)	(0.706)	(1.324)	(3.639)		

This table illustrates the results from the monthly dummy variable regressions of fund performance on the rating quintiles by using the following performance measures: the geometric mean of monthly returns and the Sharpe ratio. The performance is measured over the next 36 months after a given fund rating. Funds are rated monthly from December 2002 to December 2009. The post-rating performance is measured between December 2002 and December 2010. The category "All funds" includes all the 65 FondsNote fund rating categories. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. The *t*-statistics are adjusted using the Newey-West procedure with a lag of 35 months. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

The FondsNote has the highest predictability when using the Jensen alpha as the performance metric. Tables A2 and A3 show that better rated funds significantly differ in performance from the worst-rated funds. Jensen's alpha is frequently higher than that of the worst-rated funds. Best results are found for a three-year post-rating period. In six out of eight cases, the best-rated funds have a higher performance than the worst-rated funds. Only the best-rated funds of the equity category "World" and fixed-come category "EUR" have a significant negative performance.

These results are in line with the literature analyzing the German fund rating market (e.g. Füss et al. (2010), Müller and Weber (2014), Meinhardt (2014a), and Meinhardt (2014b)). Füss et al. (2010), for example, analyze the Morningstar rating by using the geometric mean of monthly returns and the Sharpe ratio. The study shows that the predictability of the Morningstar rating is low. Only the best-rated funds partly have a higher performance than the worst-rated funds. All other rating classes do not differ in their performances. Meinhardt (2014a) and Meinhardt (2014b), who analyze the predictability of the Feri Trust rating, the Stiftung Warentest rating, and the FondsNote for the fund rating category "Germany", obtain similar results by using the geometric mean of monthly returns to analyze the performance of fund rating classes. On the contrary, Müller and Weber (2014) study the performance differences of the Stiftung Warentest rating with the index-adjusted return, the Jensen alpha, and the Carhart four-factor alpha as the benchmark. Their results show that the rating of Stiftung Warentest can frequently distinguish between well and poorly performing funds in the future.

Further results concern the absolute performance of examined fund rating classes. The performance of the worst-rated funds is frequently positive irrespective of the performance measure and the post-rating period analyzed. Sharpe ratios are often positive but generally smaller than one indicating that funds' returns are lower than that of the risk-free rate. Moreover, the geometric mean of monthly returns of better rated funds is frequently lower than the performance of the worst-rated funds. Nonetheless, the differences in performances are not significant and, therefore, the fund rating classes do not differ from each other.

3.2 Alternative Predictors

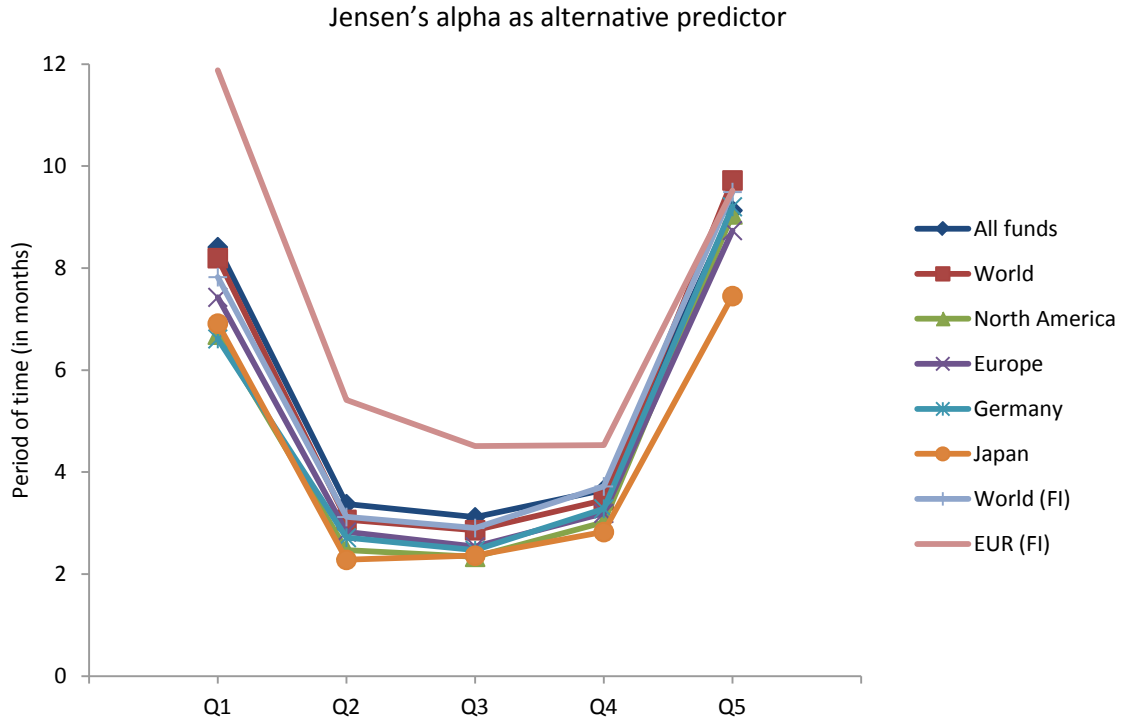
Markov chain analysis

To compare the predictability of the FondsNote with those of the alternative predictors, the Markov chain analysis and the dummy variable regression are repeated here again. The funds are ranked on the basis of four-year pre-rating performance of alternative predictors and then sorted into quintiles. Alternative predictors are the geometric mean of monthly returns, the Sharpe ratio, the index-adjusted return, and the Jensen alpha.

The period of time until the rated funds change their ratings significantly differs from that of the FondsNote. Figure 2 and Table 5 illustrate the rating persistence of the Jensen alpha as an alternative predictor. The results of the other three performance metrics are very similar to that of the Jensen alpha and can be found in Figure A1 in the Appendix.

The rating persistence of the Jensen alpha strongly varies between fund rating categories. The best results are found for the fixed-income category "EUR". The worst-rated funds especially yield the highest rating persistence. The lowest stability of fund ratings exists for the equity category "Japan". This also applies to all other alternative predictors. Compared to the FondsNote, fund ratings only persist for a short period of time. The average stability amounts to 5.31 months, which includes all fund rating categories, whereas the average stability of the FondsNote amounts to 12.55 months. This indicates that the FondsNote delivers value added in comparison to the alternative predictors, and contradicts the results of the dummy variable regressions.

Figure 2: Rating Persistence – Alternative Predictors



This figure presents the period of time (in months) until rated funds change their rating with a probability of 50%. The period of time is separately measured for different rating classes. Fund ratings are based on the four-year in-sample performance of the Jensen alpha. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0. The category "All funds" includes 65 fund rating categories evaluated by the FondsNote within the sample period from December 2002 to December 2009. The examined equity fund rating categories are: World, North America, Europe, Germany, and Japan. Fixed-income categories (FI) are: World and EUR.

Moreover, extreme fund rating classes attain the highest rating persistence (Q1 and Q5). The stability amounts up to 11.88 and 9.72 months for the worst-rated and the best-rated funds, re-

spectively. The rating persistence is the lowest for the second-worst to the second-best-rated funds irrespective of the fund rating category. These results are in line with the literature. Meinhardt (2014b) also analyzes the stability of fund ratings over time using the geometric mean of monthly returns and the Carhart four-factor alpha as an alternative predictor. The results show that the rating persistence is the highest for the extreme fund ratings. The examined fund rating stability lies between 2.51 and 10.46 months.

Table 5: Rating Persistence – Alternative Predictors

	All funds	World	North America	Europe	Germany	Japan	World (FI)	EUR (FI)
Q1	8.41	8.20	6.70	7.43	6.61	6.91	7.82	11.88
Q2	3.37	3.06	2.47	2.82	2.71	2.28	3.12	5.41
Q3	3.11	2.86	2.34	2.54	2.47	2.36	2.90	4.51
Q4	3.65	3.45	3.02	3.19	3.28	2.82	3.72	4.53
Q5	9.13	9.72	9.05	8.74	9.20	7.45	9.49	9.52

This table presents the period of time (in months) until the rated funds change their rating with a probability of 50%. Fund ratings are based on the Jensen alpha as alternative predictor. The period of time is separately measured for different rating classes. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0. The category "All funds" includes all 65 fund rating categories valuated by the FondsNote within the sample period from December 2002 to December 2009. The examined equity fund rating categories are: World, North America, Europe, Germany, and Japan. Fixed-income categories (FI) are: World and EUR.

Dummy variable regression

The methodology is identical to the evaluation of the FondsNote predictability. For the sake of brevity, the pre-rating and the post-rating performances are measured with the same performance metric. Moreover, the presented results are very similar to those received from using different performance metrics for the pre- and post-rating periods.

Tables 6, 7, A4 and A5 present the results of the dummy variable regressions. All the tables show that the predictability of the alternative predictors is similar to that of the FondsNote. Like the FondsNote, there is some support that the best-rated funds frequently have a higher performance than the worst-rated funds. Especially the three-year post-rating performance of funds, which are rated with the index-adjusted return and the Jensen alpha, differs significantly among the fund rating classes. Moreover, predictability of alternative predictors depends on the fund rating category analyzed. For example, the performance of the better rated funds frequently differs from the worst-rated funds when analyzing the fixed-income category "World". Compared to that, regression coefficients of the equity category "North America" are mainly insignificant or significantly negative. The results also show that the absolute performance is almost positive irrespective of the performance metric, the post-rating period, and the fund rating category.

Table 6:
Performance Predictability of Alternative Predictors by Using the Geometric Mean and the Sharpe Ratio as Performance Metrics over the next 12 Months

Fund sample	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
Panel A: Geometric mean of monthly returns over next 12 months							
All funds	**0.0944****	-0.0053	-0.0250	-0.0622	-0.0426	0.02	1.939
	(2.396)	(-0.098)	(-0.391)	(-1.042)	(-0.714)		
World	**0.0667***	-0.0215	-0.0180	-0.0086	0.0063	0.00	0.348
	(1.663)	(-0.350)	(-0.262)	(-0.131)	(0.083)		
North America	0.0293	-0.0055	-0.0022	0.0004	0.0092	0.00	0.081
	(0.623)	(-0.087)	(-0.033)	(0.005)	(0.127)		
Europe	0.0747	-0.0083	-0.0026	0.0025	0.0097	0.00	0.080
	(1.190)	(-0.101)	(-0.028)	(0.026)	(0.092)		
Germany	0.0308	-0.0061	0.0024	-0.0026	-0.0053	0.00	0.023
	(0.521)	(-0.069)	(0.026)	(-0.027)	(-0.050)		
Japan	**0.1202***	-0.0183	-0.0069	-0.0117	-0.0061	0.00	0.070
	(1.887)	(-0.225)	(-0.075)	(-0.115)	(-0.055)		
World (FI)	**0.0269*****	0.0026	0.0013	0.0057	0.0022	0.00	0.437
	(2.715)	(0.204)	(0.101)	(0.424)	(0.168)		
EUR (FI)	0.0248	-0.0007	0.0021	0.0023	0.0029	0.00	0.072
	(1.153)	(-0.022)	(0.070)	(0.081)	(0.105)		
Panel B: Sharpe ratio over next 12 months							
All funds	0.1669	0.0514	0.0765	0.0542	**0.3363***	0.10	**11.985*****
	(1.642)	(0.380)	(0.528)	(0.410)	(1.898)		
World	0.1283	0.0336	0.0511	0.0692	0.1027	0.01	1.059
	(1.310)	(0.241)	(0.362)	(0.466)	(0.654)		
North America	0.0791	0.0011	0.0121	0.0194	0.0299	0.00	0.145
	(1.041)	(0.010)	(0.103)	(0.159)	(0.238)		
Europe	0.2070	0.0262	0.0254	0.0301	0.0808	0.00	0.465
	(1.627)	(0.148)	(0.132)	(0.154)	(0.372)		
Germany	-0.0169	0.0075	0.0219	-0.0036	0.0160	0.00	0.062
	(-0.121)	(0.039)	(0.107)	(-0.018)	(0.077)		
Japan	**0.2376****	0.0170	0.0230	0.0248	0.0189	0.00	0.064
	(2.415)	(0.120)	(0.160)	(0.162)	(0.115)		
World (FI)	**0.2577*****	0.0300	0.0356	0.0772	**0.2870*****	0.11	**12.337*****
	(4.088)	(0.300)	(0.379)	(0.723)	(3.301)		
EUR (FI)	0.0767	0.0007	0.0175	0.0584	**0.1909***	0.07	**7.966*****
	(0.853)	(0.006)	(0.139)	(0.523)	(1.867)		

This table illustrates the results from the monthly dummy variable regressions of fund performance on the alternative predictor rating quintiles using the following performance measures: geometric mean of monthly returns and the Sharpe ratio. The performance is measured over the next 12 months after a given fund rating. Funds are rated monthly from December 2002 to December 2009. The post-rating performance is measured between December 2002 and December 2010. The category "All funds" includes all the 65 FondsNote fund rating categories. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. The *t*-statistics are adjusted by using the Newey-West procedure with a lag of eleven months. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

Although these fund ratings are based only on quantitative backward-looking performance metrics, the predictability of alternative predictors hardly differs from that of the FondsNote. The results also suggest that the performance shows persistence over time for some fund rating categories. These fund rating categories especially apply the same degree of performance predictability as the FondsNote. The results indicate that qualitative factors hardly increase the FondsNote

predictability, thus contradicting the results of the measurements for rating persistence. The reason could be that funds are only evaluated qualitatively if the quantitative FondsNote amounts up to three. Moreover, qualitative factors can only lead to a downgrade of the fund's quantitative sub-rating. This significantly limits the influence on the fund's final fund rating.

Table 7:
Performance Predictability of Alternative Predictors by Using the Geometric Mean and the Sharpe Ratio as Performance Metrics over the next 36 Months

Fund sample	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
Panel A: Geometric mean of monthly returns over next 36 months							
All funds	**0.0596*** **	-0.0061	-0.0173	**-0.0391** **	**-0.0365*** **	0.03	**3.762*** **
	(9.873)	(-0.311)	(-0.879)	(-2.438)	(-2.894)		
World	**0.0280*** **	-0.0067	-0.0069	-0.0045	0.0090	0.01	0.577
	(4.105)	(-0.406)	(-0.449)	(-0.308)	(0.555)		
North America	0.0141	-0.0042	-0.0029	-0.0043	0.0081	0.00	0.307
	(1.379)	(-0.232)	(-0.176)	(-0.247)	(0.657)		
Europe	**0.0328*** **	-0.0062	-0.0047	-0.0030	0.0080	0.00	0.251
	(3.100)	(-0.183)	(-0.217)	(-0.104)	(0.302)		
Germany	0.0034	-0.0087	-0.0099	-0.0219	-0.0319	0.01	1.491
	(0.379)	(-0.420)	(-0.584)	(-1.145)	(-1.402)		
Japan	**0.0543*** **	-0.0055	0.0038	0.0065	0.0182	0.01	0.603
	(5.857)	(-0.239)	(0.164)	(0.230)	(0.682)		
World (FI)	**0.0151*** **	0.0031	0.0040	0.0056	**0.0064** **	0.04	**4.475*** **
	(8.136)	(0.702)	(1.062)	(1.304)	(2.096)		
EUR (FI)	**0.0160* **	0.0035	0.0062	0.0059	0.0056	0.01	0.842
	(1.945)	(0.245)	(0.638)	(0.502)	(0.587)		
Panel B: Sharpe ratio over next 36 months							
All funds	**0.1427*** **	-0.0018	-0.0037	-0.0182	**0.1390** **	0.08	**9.634*** **
	(9.286)	(-0.037)	(-0.085)	(-0.581)	(2.037)		
World	**0.0715*** **	0.0149	0.0196	0.0254	0.0564	0.01	1.247
	(5.028)	(0.342)	(0.560)	(0.725)	(1.325)		
North America	**0.0510*** **	-0.0003	0.0065	0.0115	0.0288	0.00	0.504
	(3.227)	(-0.009)	(0.228)	(0.414)	(1.230)		
Europe	**0.1114*** **	-0.0046	0.0071	0.0053	0.0377	0.00	0.502
	(5.856)	(-0.060)	(0.111)	(0.075)	(0.547)		
Germany	0.0078	-0.0225	-0.0100	-0.0394	-0.0525	0.01	1.451
	(0.500)	(-0.628)	(-0.365)	(-1.186)	(-1.511)		
Japan	**0.1149*** **	0.0198	0.0226	0.0367	0.0548	0.01	1.085
	(8.780)	(0.362)	(0.520)	(0.605)	(0.932)		
World (FI)	**0.1197*** **	**0.0360** **	**0.0521** **	**0.0668*** **	**0.1195*** **	0.25	**34.226*** **
	(14.977)	(2.288)	(2.564)	(3.068)	(7.479)		
EUR (FI)	0.0575	0.0145	0.0287	0.0501	**0.1168*** **	0.15	**18.678*** **
	(1.541)	(0.239)	(0.696)	(1.020)	(2.871)		

This table illustrates the results from the monthly dummy variable regressions of fund performance on the alternative predictor rating quintiles using the following performance measures: geometric mean of monthly returns and the Sharpe ratio. The performance is measured over the next 36 months after a given fund rating. Funds are rated monthly from December 2002 to December 2009. The post-rating performance is measured between December 2002 and December 2010. The category "All funds" includes all the 65 FondsNote fund rating categories. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. The *t*-statistics are adjusted by using the Newey-West procedure with a lag of 35 months. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

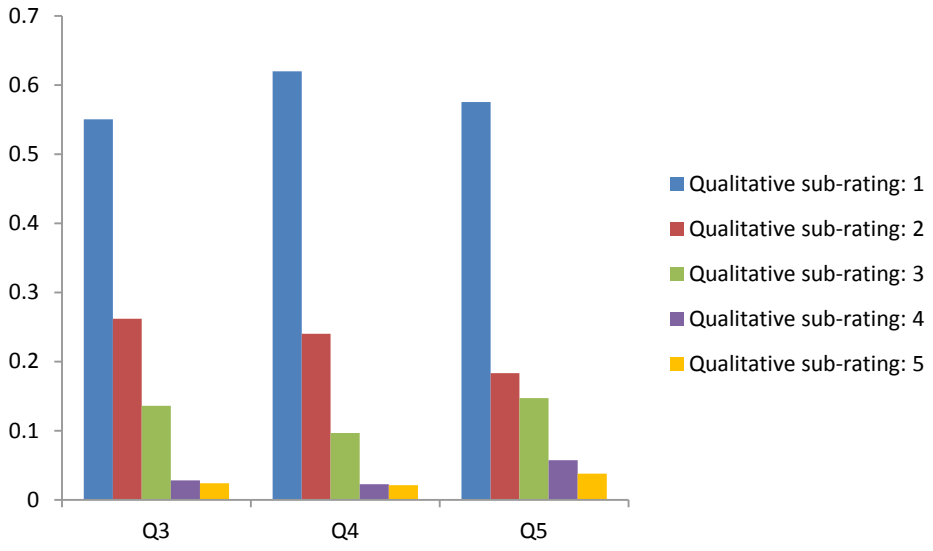
3.3 Factors Influencing the Fund Rating Predictability

Evaluating funds with quantitative and qualitative factors

Funds are only rated qualitatively if they achieve a quantitative sub-rating that lies between one and three (from quintile 3 to quintile 5). However, a qualitative sub-rating does not enhance but can diminish the final grade of a fund. The qualitative sub-rating, therefore, works as an early warning system (Axel Springer and FondsConsult (2007)). Consequently, the final grade of a fund is identical to the quantitative sub-rating or below that rating. Upgrades from the quantitative sub-rating are impossible.

Figure 3 compares the quantitative and the qualitative sub-ratings. The results show that the best qualitative rating is most frequently assigned independent of the quantitative rating. The number of funds that receive the best qualitative rating is at least two-times higher than that of the funds with the second-best qualitative rating. This indicates that the influence on the final grade is low. Funds rarely have a FondsNote that is below the quantitative sub-rating.

Figure 3: Quantitative and Qualitative Sub-Ratings



This figure presents the average probability (left axis) that a quantitatively rated fund receives the best to the worst qualitatively sub-rating in the same months within the sample period from December 2002 to December 2009. Quintile 3 (Q3) to quintile 5 (Q5) reflect funds that get the average to the best quantitative sub-rating. Analyses are based on all the funds evaluated by Axel Springer and FondsConsult and, therefore, include all 65 fund rating categories.

Table 8 illustrates the distribution of change events over fund rating classes within the sample period from December 2002 to December 2009. This enables us to gain an insight into the influence of the qualitative sub-rating on the FondsNote. Out of the 412,080 data points (4,848 funds, 85 months), we can identify 15,460 fund rating change events for the FondsNote and 13,828 for

the quantitative sub-rating. The amount of upgrades and downgrades is nearly equal. Moreover, the density of fund rating changes is bell shaped, with smaller amounts of changes toward the best and the worst rating.

Table 8 points out two important results in the context of the analysis on the influence of qualitative factors on the FondsNote. First, the probability of fund rating changes slightly increases when the qualitative factors are included in the evaluation process. The number of upgrades and downgrades are a bit lower for the quantitative sub-rating than for the FondsNote, which includes both quantitative and qualitative factors. Second, the qualitative factors reduce the fund's probability of two-class upgrades and downgrades. Two-class fund rating changes are found more often for the quantitative sub-rating than for the FondsNote. Nonetheless, one-class fund rating changes are predominant irrespective of the fund rating analyzed. The number of one-class upgrades is significantly higher for the FondsNote than for the quantitative sub-rating. This result seems to contradict the relationship that the fund's FondsNote can only be as high as the quantitative sub-rating. The solution is that two-class upgrades for the FondsNote are found less frequently. Therefore, the number of one-class upgrades increases strongly. The same is true when comparing one- and two-class downgrades.

Table 8: Change Events of the FondsNote and the Quantitative Sub-Rating

	FondsNote	Quantitative sub-rating
One-class changes	15,323	12,283
Two-class changes (upgrade)	62	687
Two-class changes (downgrade)	55	833
Three-class changes (upgrade)	11	10
Three-class changes (downgrade)	6	11
Four-class changes (upgrade)	3	4
Four-class changes (downgrade)	0	0
Total changes	15,460	13,828

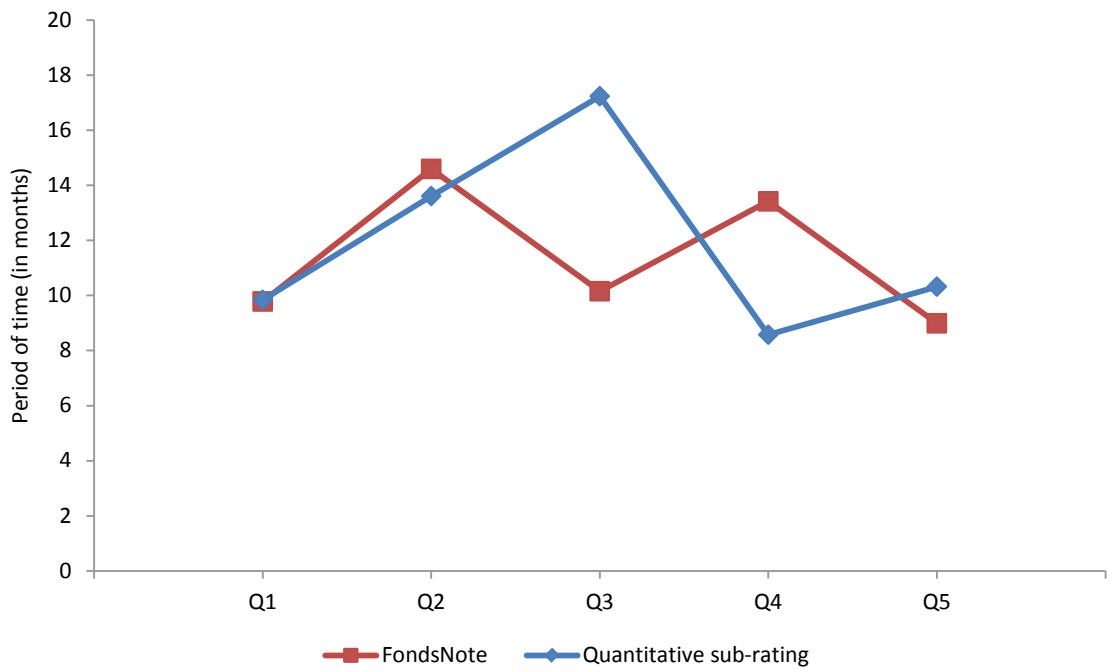
	FondsNote		Quantitative sub-rating	
Initial fund rating	One-class upgrade	One-class downgrade	One-class upgrade	One-class downgrade
1		1,279		735
2	1,008	2,983	714	2,107
3	2,567	2,922	1,525	2,781
4	2,575	996	2,485	971
5	993		965	
Total	7,143	8,180	5,689	6,594

This table presents the number of fund rating changes for funds rated with the FondsNote and the quantitative sub-rating. The analysis includes all the 65 fund rating categories within the sample period from December 2002 to December 2009.

How does this influence the rating stability over time? Change events of fund ratings are more frequently found for the FondsNote than for the quantitative sub-rating. Consequently, fund rating stability should be lower for funds rated with the FondsNote compared to the funds that receive a quantitative sub-rating only. Figure 4 presents the results of the rating persistence measure.

The FondsNote and the quantitative sub-rating only differ in fund rating stability between the average-rated (Q3) and the second-best-rated funds (Q4). Fund rating stability is much higher for funds with an average quantitative sub-rating compared to these funds rated with a FondsNote. The opposite is true for the second-best-rated funds. All other fund rating classes rarely differ in fund rating stability.

Figure 4: Rating Persistence of the Quantitative Sub-Rating



This figure presents the period of time (left axis: in months) till the rated funds change their rating with a probability of 50%. The period of time is separately measured for different rating classes. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0. The rating persistence is measured for the FondsNote and the quantitative sub-rating, including all the 65 fund rating categories within the sample period from December 2002 to December 2009.

It is hardly surprising that the differences in fund rating stability can rarely be found for worse rated funds because these funds are almost completely rated quantitatively. Interestingly, the fund rating stability of the best-rated funds (Q5) is similar between the FondsNote and the quantitative sub-rating. The reason is that these funds not only receive the best quantitative sub-rating, but they also frequently achieve the best qualitative sub-rating.

Tables 9 and 10 present the dummy variable regression results for the quantitative and qualitative sub-ratings, respectively. This can be seen as a scenario analysis for the FondsNote rating. On the one hand, the weight of the qualitative sub-rating is zero and, therefore, the FondsNote is only dependent on the quantitative sub-rating (Table 9). On the other hand, the weight of the quantitative sub-rating is zero and, therefore, the FondsNote only depends on the qualitative sub-rating (Table 10). For the sake of brevity, the results for the different weights are not shown because they hardly differ from the one presented.

Tables 9 and 10 illustrate that the future performance does not differ between fund rating classes when only using the quantitative or the qualitative sub-rating as the performance predictor. The results are very similar between both fund ratings irrespective of the post-rating period and the performance metric. The performance of fund rating classes only differs from each other when using the qualitative sub-rating as the Jensen alpha predictor for a post-rating period of 36 months (Table 10). Nonetheless, the better rated funds perform significantly worse than the lowest-rated funds.

Table 9: Dummy Variable Regression – Quantitative Sub-Rating

	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
Panel A: Quantitative sub-rating – Performance over next 12 months							
Geometric mean	0.0733	-0.0082	-0.0123	-0.0164	-0.0034	0.00	0.154
	(1.498)	(-0.139)	(-0.191)	(-0.248)	(-0.052)		
Sharpe ratio	**0.2630****	0.0491	0.1347	0.1028	0.1833	0.03	**3.683*****
	(2.266)	(0.327)	(0.747)	(0.557)	(1.229)		
Index-adjusted return	**0.0019*****	0.0002	0.0004	0.0002	0.0004	0.00	0.474
	(13.539)	(0.450)	(0.613)	(0.252)	(0.467)		
Jensen alpha	**0.0018*****	0.0000	0.0004	0.0003	**0.0012****	0.05	**5.561*****
	(5.537)	(-0.079)	(1.015)	(0.621)	(2.406)		
Panel B: Quantitative sub-rating – Performance over next 36 months							
Geometric mean	**0.0443*****	-0.0107	-0.0104	-0.0180	-0.0022	0.01	0.715
	(4.016)	(-0.493)	(-0.652)	(-1.014)	(-0.140)		
Sharpe ratio	**0.1556*****	0.0262	0.0666	0.0524	**0.0927***	0.03	**2.745****
	(6.976)	(0.435)	(1.302)	(0.781)	(1.824)		
Index-adjusted return	**0.0019*****	-0.0003	-0.0001	-0.0004	0.0001	0.01	1.522
	(6.817)	(-0.488)	(-0.169)	(-0.697)	(0.289)		
Jensen alpha	**0.0019*****	**-0.0004****	-0.0003	-0.0004	0.0004	0.09	**9.806*****
	(11.898)	(-2.395)	(-0.808)	(-1.289)	(1.391)		

This table illustrates the results from the monthly dummy variable regressions of fund performance on the rating quintiles using the following performance measures: geometric mean of monthly returns, Sharpe ratio, index-adjusted return, and Jensen alpha. The performance is measured over the next 12 and 36 months after a given fund rating. Funds are rated monthly from December 2002 to December 2009. The post-rating performance is measured between December 2002 and December 2010. The fund sample includes all the 65 FondsNote fund rating categories. Fund rating quintiles are solely based on the quantitative sub-rating excluding the qualitative sub-rating. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. The *t*-statistics are adjusted by using the Newey-West procedure with a lag of 11 or 35 months. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

The level of predictability barely differs between the sub-ratings and the FondsNote because the influence of the qualitative rating on the quantitative rating is low. Therefore, adjustments of sub-rating weights within the scope of optimization techniques do not increase the predictability of the FondsNote.

There are various reasons for the low predictability. Funds can only diminish their quantitative sub-rating when they receive a qualitative sub-rating. Moreover, only the best to the average-rated funds are evaluated qualitatively. These funds frequently maintain the best qualitative sub-rating (see Figure 3). Furthermore, lower rated funds solely receive the quantitative sub-rating as final grade. The analyses also reveal that a fund's qualitative sub-rating barely changes over time. This is not surprising since the qualitative sub-rating depends on factors like the investment approach and the transparency of the fund, which will almost remain unchanged. However, the continuity of the management only has a weight of 40% within the overall qualitative sub-rating.

Table 10: Dummy Variable Regression – Qualitative Sub-Rating

	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
Panel A: Qualitative sub-rating – Performance over next 12 months							
Geometric mean	**0.0663***	0.0032	-0.0044	-0.0034	-0.0062	0.00	0.055
	(1.673)	(0.057)	(-0.078)	(-0.060)	(-0.112)		
Sharpe ratio	**0.2485*****	0.0102	0.2308	0.1367	0.1760	0.06	**6.835*****
	(2.944)	(0.086)	(1.163)	(1.040)	(1.147)		
Index-adjusted return	**0.0025****	-0.0002	-0.0003	0.0000	0.0001	0.00	0.257
	(2.202)	(-0.173)	(-0.239)	(-0.031)	(0.041)		
Jensen alpha	**0.0032*****	-0.0010	-0.0009	-0.0006	-0.0005	0.02	**2.497****
	(4.199)	(-1.255)	(-1.074)	(-0.781)	(-0.681)		
Panel B: Qualitative sub-rating – Performance over next 36 months							
Geometric mean	**0.0476*****	-0.0077	-0.0097	-0.0065	-0.0109	0.00	0.311
	(6.151)	(-0.533)	(-0.662)	(-0.453)	(-0.882)		
Sharpe ratio	**0.1676*****	-0.0030	0.0945	0.0337	0.0603	0.03	**3.663*****
	(15.249)	(-0.087)	(1.087)	(0.893)	(1.013)		
Index-adjusted return	**0.0025*****	-0.0003	-0.0005	-0.0005	-0.0004	0.01	1.323
	(6.627)	(-0.482)	(-1.074)	(-1.009)	(-0.948)		
Jensen alpha	**0.0029*****	**-0.0009*****	**-0.0012*****	**-0.0010*****	**-0.0009*****	0.11	**13.264*****
	(13.366)	(-2.588)	(-4.990)	(-3.841)	(-3.159)		

This table illustrates the results from the monthly dummy variable regressions of fund performance on the rating quintiles using the following performance measures: geometric mean of monthly returns, Sharpe ratio, index-adjusted return, and Jensen alpha. The performance is measured over the next 12 and 36 months after a given fund rating. Funds are rated monthly from December 2002 to December 2009. The post-rating performance is measured between December 2002 and December 2010. The fund sample includes all the 65 FondsNote fund rating categories. Fund rating quintiles are solely based on the qualitative sub-rating, excluding the quantitative sub-rating. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. The *t*-statistics are adjusted by using the Newey-West procedure with a lag of 11 or 35 months. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

Qualitative factors that are more sensitive to changes would lead to a higher difference in predictability between the qualitative and the quantitative sub-ratings. For instance, literature shows

that the activity of the fund management, also called Active Share, plays an important role in performance persistence. Funds with portfolio holdings that frequently differ from their benchmark holdings exhibit strong performance persistence (e.g. Cremers and Petajisto (2009) and Petajisto (2013)).

Costs as an influencing factor

This section answers the question of whether the costs of funds are responsible for the low predictability of the FondsNote and the alternative predictors. Table 11 presents the average total expense ratio for the entire sample period and points out that the worst-rated funds (Q1) frequently possess the highest costs. Values range between 1.15% per year for the fixed-income category "EUR" and 1.93% per year for the equity category "North America". The lowest TERs are found for quintile 3 and 4, which represent the average-rated and second-best-rated funds. Moreover, the best-rated funds (Q5) frequently have lower costs than the worst-rated funds. The *t*-values are often negative and significantly different from zero. These results are also confirmed by Figure A2 in the Appendix, which shows the average total expense ratio of the rating quintiles over time.

Table 11: Analysis of Total Expense Ratios

	Q1	Q2	Q3	Q4	Q5	Q5-Q1	T-Stat.
Panel A: TER of the original fund rating							
FondsNote	1.70	1.54	1.44	1.42	1.52	**-0.17*** **	(-21.783)
Quantitative sub-rating	1.70	1.54	1.43	1.44	1.50	**-0.20*** **	(-26.797)
Qualitative sub-rating	1.57	1.49	1.44	1.48	1.41	**-0.17*** **	(-12.233)
Panel B: TER of alternative predictors							
Geometric mean of monthly returns	1.51	1.48	1.53	1.44	1.58	**0.07*** **	(14.154)
Sharpe ratio	1.58	1.59	1.58	1.59	1.23	**-0.36*** **	(-32.075)
Index-adjusted return	1.75	1.63	1.59	1.30	1.38	**-0.38*** **	(-34.297)
Jensen's alpha	1.73	1.59	1.50	1.38	1.45	**-0.28*** **	(-23.396)
Panel C: TER of fund rating categories							
World	1.83	1.70	1.68	1.65	1.80	-0.03	(-0.950)
North America	1.93	1.88	1.73	1.77	1.97	0.04	(1.210)
Europe	1.81	1.67	1.64	1.70	1.88	**0.07*** **	(3.054)
Germany	1.76	1.41	1.32	1.43	1.55	**-0.22*** **	(-5.724)
Japan	1.82	1.77	1.70	1.65	1.81	-0.01	(-0.267)
World (FI)	1.27	1.11	1.14	1.07	0.99	**-0.27*** **	(-14.059)
EUR (FI)	1.15	1.02	0.88	0.78	0.79	**-0.36*** **	(-13.960)

This table presents the relationship between fund rating classes and costs. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund ratings. Costs are defined as the total expense ratio of funds per year (arithmetic mean of fund rating classes) and are presented in percentages. The sample period comprises all the funds rated from December 2002 to December 2009. TERs in Panel A and B are applied for the entire fund universe including all the 65 different fund rating categories. Panel C presents TERs of the fund rating classes for the five equity categories: World, North America, Europe, Germany, and Japan. Moreover, two fixed-income categories (FI) are analyzed: World (FI), and EUR (FI). ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

Summing up, the costs of funds only partially influence the FondsNote predictability and the alternative predictors. For instance, as shown in Table 2, the best-rated funds outperform the worst-rated funds by 1.68% per year considering the category "All funds" and the Jensen alpha over the next 12 months which is substantially higher than the 0.17% difference in TERs. Nonetheless, the FondsNote predictability would still be lower if the differences in TERs between fund rating classes were smaller.

The behavior of investors as an influencing factor

Following the analyses on the influence of qualitative factors and costs on fund rating predictability, this section will detail the relationship between fund ratings and investment behavior. Table 12 presents the relationship between fund rating classes and monthly fund flows.

Table 12: Analysis of Fund Flows

	Q1	Q2	Q3	Q4	Q5	Q5-Q1	T-Stat.
Panel A: TER of the original fund rating							
FondsNote	-0.90	-2.28	-2.47	0.40	4.90	**5.80*** **	(4.299)
FondsNote - Bull market	-0.98	-2.54	-2.66	1.90	6.61	**7.59*** **	(3.592)
FondsNote - Bear market	-2.14	-3.10	-3.95	-3.54	-1.03	1.11	(0.393)
Quantitative sub-rating	-0.89	-2.30	-1.94	1.08	3.33	**4.22*** **	(3.473)
Qualitative sub-rating	-0.23	2.53	-0.71	-0.56	-0.98	-0.75	(-0.203)
Panel B: TER of alternative predictors							
Geometric mean of monthly returns	-2.71	-2.09	-0.08	-1.97	2.08	**4.79*** **	(7.171)
Sharpe ratio	-2.22	-1.63	-0.40	0.36	-0.93	**1.29***	(1.853)
Index-adjusted return	-2.26	-2.49	-0.02	-0.98	2.20	**4.47*** **	(8.675)
Jensen's alpha	-2.68	-2.26	-1.32	-0.74	3.40	**6.08*** **	(10.176)
Panel C: TER of fund rating categories							
World	-0.33	-2.46	-1.37	1.84	5.50	**5.83*** **	(2.675)
North America	-1.44	-1.64	-1.15	1.86	0.50	1.94	(1.443)
Europe	-0.37	-3.00	-2.53	-0.09	4.60	**4.97*** **	(4.469)
Germany	-0.47	-1.32	-1.99	0.34	0.75	1.21	(1.116)
Japan	-0.22	-1.30	0.49	0.67	-15.03	**-14.82***	(-1.786)
World (FI)	-1.04	-1.22	-1.84	0.69	4.79	**5.83***	(1.651)
EUR (FI)	-0.87	-4.39	-3.57	-0.30	4.51	**5.38** **	(2.175)

This table presents the relationship between fund rating classes and fund flows. Quintile 1 (Q1) to quintile 5 (Q5) reflect the funds that get the worst to the best fund rating. Fund flows are presented in absolute terms (in millions of euros) on a monthly basis. The sample period comprises all the rated funds from December 2002 to December 2009. Moreover, two sub-periods are analyzed: the bull market (December 2002 to June 2007) and the bear market (December 2007 to February 2009). Fund flows in Panel A and B are applied for the entire fund universe including all the 65 different fund rating categories. Panel C presents fund flows of fund rating classes for five equity categories: World, North America, Europe, Germany, and Japan. Moreover, two fixed-income categories (FI) are analyzed: World (FI), and EUR (FI). ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

Inspections of the results reveal that the best-rated funds (Q5) frequently have the highest inflows compared to all the other fund rating classes. Funds in quintile 1 to 4 receive slightly negative fund flows. Moreover, fund flows of the best-rated funds significantly differ from the worst-

rated funds and are mainly positive. Only some fund rating categories have insignificant fund flow differences.

In line with Füss et al. (2010) and Meinhardt (2014a), the investment behavior differs between a bullish and bearish capital market environment. Inflows to the best-rated funds are the highest for the bullish capital market and considerably lower for the bearish capital market. Moreover, all fund rating classes are faced with outflows in the bearish capital market environment. Figure A3 in the Appendix also illustrates these results.

The assumption that fund ratings influence the behavior of investors can be confirmed by the results presented. The best-rated funds especially receive disproportionately larger inflows. However, lower rated funds barely face cash outflows. Fund rating predictability is significantly influenced as long as fund flows differ among the fund rating classes.

4 Conclusion

The FondsNote predictability primarily depends on the performance metric used and it significantly differs among the fund rating categories. Predictability is the highest when using the Jensen alpha for a three-year post-rating period. However, fund rating classes barely differ in performance when using the geometric mean of monthly returns and the Sharpe ratio as performance measures. Surprisingly, the fund rating stability over time is higher compared to other fund ratings provided by the Morningstar, Feri Trust, and Stiftung Warentest ratings. Fund ratings based on alternative predictors are similarly poor in forecasting the future performance of funds. The stability of these fund ratings is even lower than that of the FondsNote.

This study discusses several factors that can influence the FondsNote predictability. The results indicate that qualitative factors, which are used in the evaluation process, barely hike the overall predictability of the fund rating. This result is confirmed by a detailed analysis of both the quantitative and the qualitative sub-ratings. Irrespective of the sub-rating weight used in the FondsNote, the differences in performance are frequently insignificant between the different fund rating classes. Qualitative factors used by Axel Springer and FondsConsult are apparently unable to create additional value if they are not supplemented by stronger time-varying factors. These could be factors such as management turnover, style-consistency and the fund's activity to significantly differ between portfolio holdings and index holdings.

Moreover, this study investigates the influence of costs as well as the behavior of investors on the FondsNote ratings. The results show that the costs vary significantly between the fund rating classes and they are, therefore, a partial cause for the low predictability. Just like in the case of

costs of funds, the fund flows differ among the fund rating classes. The best-rated funds especially obtain disproportionately larger inflows. The worse rated funds frequently face outflows. Thus, the FondsNote predictability is influenced by the behavior of investors.

Future research may investigate optimization techniques for further qualitative factors to improve the predictability of fund ratings. Fund rating agencies such as Feri Trust, Standard & Poor's, Sauren, and Scope use similar qualitative factors to those that are used in the FondsNote evaluation process. This could indicate that the predictability of their fund ratings is also low. To investigate this topic, the predictability of further fund rating agencies has to be analyzed in detail.

References

Antypas, A., Caporale, G. M., Kourogenis, N., Pittas, N. (2009). Selectivity, market timing and the Morningstar star-rating system. Working paper, No. 874, DIW Berlin, April.

Axel Springer, FondsConsult (2007). FondsNote. Das Fonds-Rating von Axel Springer Finanzen Verlag und FondsConsult. Axel Springer Finanzen Verlag & FondsConsult Research AG, Munich.

Baker, H. K., Haslem, J. A., Smith, D. M. (2009). Performance and characteristics of actively managed institutional equity mutual funds. The Journal of Investing 18 (1), pp. 27-44.

Barelkowsa, M. (2010). Management turnover and fund performance: evidence from German mutual funds. Working paper, Humboldt University of Berlin, November.

Bessler, W., Blake, D. P., Lückoff, P., Tonks, I. (2010). Why does mutual fund performance not persist? The impact and interaction of fund flows and manager changes. Working paper, presented at the Finance Meeting EUROFIDAI – AFFI in Paris in December 2010.

Blake, C. R., Morey, M. R. (2000). Morningstar ratings and mutual fund performance. The Journal of Financial and Quantitative Analysis 35 (3), pp. 451-483.

Brown, K. C., Harlow, W. V., Zhang, H. (2009). Staying the course: the role of investment style consistency in the performance of mutual funds. Working paper, University of Texas, March.

BVI (2007a). Was sagen Rankings und Ratings aus? BVI Jahrbuch 2007, Bundesverband Investment und Asset Management e.V., pp. 65-71.

BVI (2007b). Ranking-/Rating-Transparenz-Standard – Antwortsammlung. Bundesverband Investment und Asset Management e.V., 3rd edition, 10/2007.

Carhart, M. (1997). On the persistence of mutual fund performance. The Journal of Finance 52 (1), pp. 57-82.

Chen, J., Hong, H., Huang, M., Kubik, J. D. (2004). Does fund size erode mutual fund performance? The role of liquidity and organization. The American Economic Review 94 (5), pp. 1276-1302.

Cremers, K. J. M., Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. The Review of Financial Studies 22 (9), pp. 3329-3365.

Del Guercio, D., Tkac, P. A. (2008). Star power: the effect of Morningstar ratings on mutual fund flow. The Journal of Financial and Quantitative Analysis 43 (4), pp. 907-936.

Duret, A., Hereil, P., Mitaine, P., Moussavi, N., Roncalli, T. (2008). Fund rating systems and performance predictability. Working paper, University of Evry, April.

Elton, E. J., Gruber, M. J., Das, S., Hlavka, M. (1993). Efficiency with costly information: a reinterpretation of evidence from managed portfolios. The Review of Financial Studies 6 (1), pp. 1-22.

Fama, E. F., MacBeth, J. D. (1973). Risk, return, and equilibrium: empirical tests. The Journal of Political Economy 81 (3), pp. 607-636.

Füss, R., Hille, J., Rindler, P., Schmidt, J., Schmidt, M. (2010). From rising stars and falling angels: on the relationship between the performance and ratings of German mutual funds. Journal of Wealth Management 13 (1), pp. 75-90.

Garnier, O., Pujol, T. (2007). Can today's stars be used to read the stars of the future? Working paper, No. 3, March.

Gerrans, P. (2006). Morningstar ratings and future performance. Accounting & Finance 46 (4), pp. 605-628.

Hereil, P., Mitaine, P., Moussavi, N., Roncalli, T. (2010). Mutual fund ratings & performance persistence. White paper 3, Quant Research by Lyxor, June.

Huij, J., Lansdorp, S., Verbeek, M. (2012). Managerial turnover and the behavior of mutual fund investors. Working paper, Rotterdam School of Management, October.

Israel, R. B., Rosenthal, J. S., Wei, J. Z. (2001). Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings. Mathematical Finance 11 (2), pp. 245-265.

Kräussl, R., Sandelowsky, R. (2007). The predictive performance of Morningstar's mutual fund ratings. Working paper, VU University Amsterdam, August.

Meinhardt, C. (2014a). Fondsbewertungen in Deutschland – Ein Vergleich der Vorhersagekraft. In: dissertation: essays on actively and passively managed financial products, Humboldt University of Berlin, pp. 39-71, May 2015.

Meinhardt, C. (2014b). Ratings and performance of German mutual funds – A comparison of Feri Trust, Finanztest and FondsNote. In: dissertation: essays on actively and passively managed financial products, Humboldt University of Berlin, pp. 72-101, May 2015.

Morey, M. R., Gottesman, A. (2006). Morningstar mutual fund ratings redux. Journal of Investment Consulting 8 (1), pp. 25-37.

Müller, S., Weber, M. (2014). Evaluating the rating of Stiftung Warentest: how good are mutual fund ratings and can they be improved? European Financial Management 20 (2), pp. 207-235.

Petajisto, A. (2013). Active share and mutual fund performance. Financial Analysts Journal 69 (4), pp. 73-93.

Sirri, E. R., Tufano, P. (1998). Costly search and mutual fund flows, The Journal of Finance 53 (5), pp. 1589-1622.

Appendix

Table A1: Monthly Transition Matrix

FondsNote	Q5	Q4	Q3	Q2	Q1	NR
Q5	0.9202	0.0747	0.0009	0.0001	0	0.0041
Q4	0.0158	0.9318	0.0467	0.0005	0.0001	0.0052
Q3	0.0003	0.0364	0.9167	0.0414	0.0001	0.0050
Q2	0.0001	0.0006	0.0381	0.9371	0.0147	0.0094
Q1	0.0002	0.0003	0.0003	0.0573	0.9270	0.0148
NR	0.0038	0.0084	0.0076	0.0075	0.0031	0.9696

This table presents the empirical monthly transition matrix based on the FondsNote. The fund sample includes all the funds rated by Axel Springer and FondsConsult (funds of 65 fund rating categories) within the sample period from December 2002 to December 2009. The best-rated to the worst-rated funds are represented by quintile 5 (Q5) to 1 (Q1). The class "NR" states unrated funds that are included or excluded from the universe of rated funds due to liquidation, mergers or changes in investment focus.

Table A2:
Performance Predictability of the FondsNote by Using the Index-Adjusted Return and the Jensen Alpha as Performance Metrics over the next 12 Months

Fund sample	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
Panel A: Index-adjusted return over next 12 months							
All funds	**0.0019*** **	0.0003	0.0002	0.0005	0.0003	0.01	0.624
	(13.635)	(0.471)	(0.381)	(0.687)	(0.318)		
World	-0.0004	0.0022	**0.0015***	**0.0025** **	0.0022	0.04	**4.827*** **
	(-0.770)	(1.451)	(1.764)	(2.564)	(0.895)		
North America	-0.0001	0.0001	0.0003	0.0003	**0.0021*** **	0.09	**10.773*** **
	(-0.212)	(0.159)	(0.356)	(0.336)	(2.664)		
Europe	**0.0013***	-0.0006	-0.0007	0.0002	0.0002	0.02	**2.690** **
	(1.935)	(-0.788)	(-0.963)	(0.198)	(0.301)		
Germany	0.0003	-0.0002	-0.0003	-0.0004	0.0006	0.01	1.444
	(0.291)	(-0.152)	(-0.231)	(-0.271)	(0.342)		
Japan	0.0018	-0.0004	-0.0007	-0.0005	-0.0010	0.01	1.138
	(1.584)	(-0.330)	(-0.517)	(-0.410)	(-0.785)		
World (FI)	**0.0045*** **	-0.0009	-0.0005	-0.0005	-0.0005	0.02	**2.377***
	(9.068)	(-1.647)	(-0.967)	(-0.729)	(-0.334)		
EUR (FI)	**0.0060*** **	-0.0007	-0.0006	-0.0009	-0.0008	0.00	0.301
	(8.791)	(-0.595)	(-0.471)	(-0.641)	(-0.455)		
Panel B: Jensen's alpha over next 12 months							
All funds	**0.0018*** **	0.0000	0.0003	0.0006	**0.0014** **	0.06	**6.916*** **
	(5.475)	(-0.064)	(0.621)	(1.316)	(2.488)		
World	**-0.0016*** **	0.0030	**0.0027*** **	**0.0037*** **	**0.0044*** **	0.12	**14.174*** **
	(-3.089)	(1.617)	(3.096)	(4.133)	(4.328)		
North America	-0.0011	0.0006	0.0009	0.0013	**0.0040*** **	0.20	**27.061*** **
	(-0.878)	(0.407)	(0.687)	(1.004)	(2.841)		
Europe	-0.0006	0.0005	0.0010	0.0017	**0.0037***	0.12	**14.594*** **
	(-0.293)	(0.217)	(0.474)	(0.874)	(1.721)		
Germany	-0.0001	0.0000	-0.0001	-0.0001	0.0010	0.02	**2.205***
	(-0.097)	(-0.013)	(-0.058)	(-0.054)	(0.617)		
Japan	0.0024	-0.0002	-0.0004	-0.0002	0.0013	0.02	1.730
	(1.367)	(-0.109)	(-0.202)	(-0.103)	(0.532)		
World (FI)	**0.0035*** **	-0.0004	0.0001	0.0003	-0.0003	0.02	1.950
	(6.031)	(-0.625)	(0.222)	(0.403)	(-0.239)		
EUR (FI)	**0.0046*** **	**-0.0015***	**-0.0018***	**-0.0017***	-0.0012	0.04	**4.365*** **
	(6.204)	(-1.872)	(-1.893)	(-1.773)	(-1.316)		

This table presents the results from the monthly dummy variable regressions of fund performance on the rating quintiles using the following performance measures: the index-adjusted return and the Jensen alpha. The performance is measured over the next 12 months after a given fund rating. Funds are rated monthly from December 2002 to December 2009. The post-rating performance is measured between December 2002 and December 2010. The category "All funds" includes all the 65 FondsNote fund rating categories. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. The *t*-statistics are adjusted by using the Newey-West procedure with a lag of 11 months. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

Table A3:
Performance Predictability of the FondsNote by Using the Index-Adjusted Return and the Jensen Alpha as Performance Metrics over the next 36 Months

Fund sample	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
Panel A: Index-adjusted return over next 36 months							
All funds	**0.0019*** **	-0.0003	-0.0002	-0.0001	0.0001	0.01	0.831
	(6.773)	(-0.486)	(-0.337)	(-0.165)	(0.209)		
World	**0.0062*** **	**-0.0055*** **	**-0.0057*** **	**-0.0050*** **	**-0.0042* **	0.12	**13.844*** **
	(3.952)	(-3.457)	(-3.523)	(-3.013)	(-1.924)		
North America	**0.0010*** **	**-0.0013*** **	**-0.0013*** **	**-0.0012*** **	**0.0008* **	0.24	**32.444*** **
	(5.227)	(-4.059)	(-4.480)	(-3.017)	(1.824)		
Europe	**0.0010*** **	**-0.0005* **	**-0.0003* **	-0.0001	**0.0005*** **	0.11	**12.389*** **
	(14.618)	(-1.652)	(-1.930)	(-0.378)	(3.040)		
Germany	**-0.0002*** **	**0.0003** **	0.0001	-0.0002	0.0002	0.04	**3.827*** **
	(-4.818)	(2.164)	(0.954)	(-1.095)	(1.251)		
Japan	**0.0005*** **	0.0001	**0.0006*** **	**0.0007*** **	**0.0006** **	0.05	**5.020*** **
	(7.080)	(0.550)	(4.083)	(3.738)	(2.508)		
World (FI)	**0.0019*** **	**0.0003* **	**0.0011*** **	**0.0008*** **	**0.0013* **	0.24	**32.832*** **
	(12.854)	(1.728)	(5.153)	(3.660)	(1.835)		
EUR (FI)	**0.0038*** **	-0.0004	-0.0001	-0.0006	-0.0003	0.02	1.715
	(24.967)	(-0.870)	(-0.242)	(-0.889)	(-0.542)		
Panel B: Jensen's alpha over next 36 months							
All funds	**0.0019*** **	**-0.0004** **	-0.0004	-0.0001	**0.0005* **	0.10	**11.206*** **
	(11.767)	(-2.410)	(-1.296)	(-0.379)	(1.746)		
World	**0.0057*** **	**-0.0051*** **	**-0.0053*** **	**-0.0045*** **	**-0.0034** **	0.09	**10.740*** **
	(3.634)	(-3.304)	(-3.292)	(-2.750)	(-2.102)		
North America	0.0001	-0.0006	**-0.0006* **	-0.0002	**0.0018*** **	0.22	**29.242*** **
	(0.277)	(-1.604)	(-1.819)	(-0.606)	(3.759)		
Europe	0.0000	0.0000	0.0003	0.0007	**0.0019*** **	0.24	**33.133*** **
	(0.093)	(-0.070)	(0.616)	(1.451)	(3.466)		
Germany	**-0.0007*** **	**0.0006** **	**0.0004** **	0.0003	**0.0005** **	0.05	**5.128*** **
	(-4.853)	(2.382)	(2.506)	(1.065)	(2.159)		
Japan	**0.0004*** **	**0.0004* **	**0.0008*** **	**0.0011*** **	**0.0020*** **	0.14	**17.037*** **
	(2.771)	(1.677)	(3.265)	(5.972)	(6.542)		
World (FI)	**0.0017*** **	0.0002	**0.0010*** **	**0.0008*** **	**0.0011*** **	0.28	**40.152*** **
	(13.220)	(1.477)	(6.581)	(4.015)	(4.487)		
EUR (FI)	**0.0033*** **	**-0.0009* **	**-0.0009* **	**-0.0011** **	**-0.0010* **	0.10	**11.895*** **
	(7.867)	(-1.926)	(-1.825)	(-2.301)	(-1.884)		

This table presents the results from the monthly dummy variable regressions of fund performance on the rating quintiles using the following performance measures: the index-adjusted return and the Jensen alpha. The performance is measured over the next 36 months after a given fund rating. Funds are rated monthly from December 2002 to December 2009. The post-rating performance is measured between December 2002 and December 2010. The category "All funds" includes all the 65 FondsNote fund rating categories. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. The *t*-statistics are adjusted by using the Newey-West procedure with a lag of 35 months. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

Table A4:
Performance Predictability of Alternative Predictors by Using the Index-Adjusted Return and the Jensen Alpha as Performance Metrics over the next 12 Months

Fund sample	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
Panel A: Index-adjusted return over next 12 months							
All funds	**0.0009****	0.0001	**0.0011***	**0.0026*****	0.0023	0.17	**21.383*****
	(2.189)	(0.119)	(1.729)	(3.135)	(1.521)		
World	0.0026	-0.0022	-0.0016	-0.0009	0.0000	0.05	**5.684*****
	(1.577)	(-1.289)	(-0.930)	(-0.488)	(0.005)		
North America	0.0000	-0.0004	0.0000	0.0002	**0.0013***	0.07	**7.774*****
	(-0.013)	(-0.607)	(-0.044)	(0.314)	(1.840)		
Europe	0.0004	-0.0001	0.0005	0.0007	0.0010	0.04	**3.822*****
	(0.881)	(-0.231)	(0.764)	(0.915)	(0.779)		
Germany	0.0007	**-0.0015***	-0.0010	-0.0015	-0.0017	0.03	**3.266****
	(1.078)	(-1.939)	(-1.110)	(-1.361)	(-0.754)		
Japan	**0.0017****	-0.0006	-0.0005	-0.0002	-0.0006	0.01	0.810
	(2.033)	(-0.643)	(-0.443)	(-0.211)	(-0.520)		
World (FI)	**0.0040*****	0.0001	0.0001	0.0007	0.0003	0.03	**3.172****
	(13.488)	(0.265)	(0.244)	(1.370)	(0.413)		
EUR (FI)	**0.0055*****	-0.0001	0.0002	-0.0002	0.0001	0.00	0.055
	(7.309)	(-0.056)	(0.138)	(-0.115)	(0.068)		
Panel B: Jensen's alpha over next 12 months							
All funds	0.0010	0.0002	0.0011	**0.0019****	**0.0021****	0.14	**16.907*****
	(1.514)	(0.350)	(1.631)	(2.306)	(2.180)		
World	0.0012	-0.0001	-0.0001	0.0007	0.0015	0.02	**2.230***
	(0.718)	(-0.060)	(-0.077)	(0.395)	(0.761)		
North America	-0.0004	-0.0003	0.0003	0.0003	0.0015	0.07	**8.112*****
	(-0.341)	(-0.209)	(0.214)	(0.271)	(1.307)		
Europe	-0.0005	0.0007	0.0007	0.0010	**0.0024****	0.09	**10.416*****
	(-0.613)	(0.565)	(0.526)	(0.823)	(2.103)		
Germany	-0.0003	-0.0010	-0.0005	-0.0008	-0.0005	0.01	0.652
	(-0.516)	(-0.990)	(-0.588)	(-0.683)	(-0.172)		
Japan	**0.0026****	-0.0002	-0.0002	-0.0005	-0.0001	0.00	0.269
	(2.115)	(-0.166)	(-0.151)	(-0.358)	(-0.078)		
World (FI)	**0.0033*****	0.0002	0.0005	**0.0010****	**0.0015*****	0.18	**23.034*****
	(11.496)	(0.510)	(1.338)	(2.047)	(3.038)		
EUR (FI)	**0.0027*****	0.0007	0.0003	0.0008	0.0013	0.04	**4.108*****
	(3.936)	(0.722)	(0.383)	(0.980)	(1.576)		

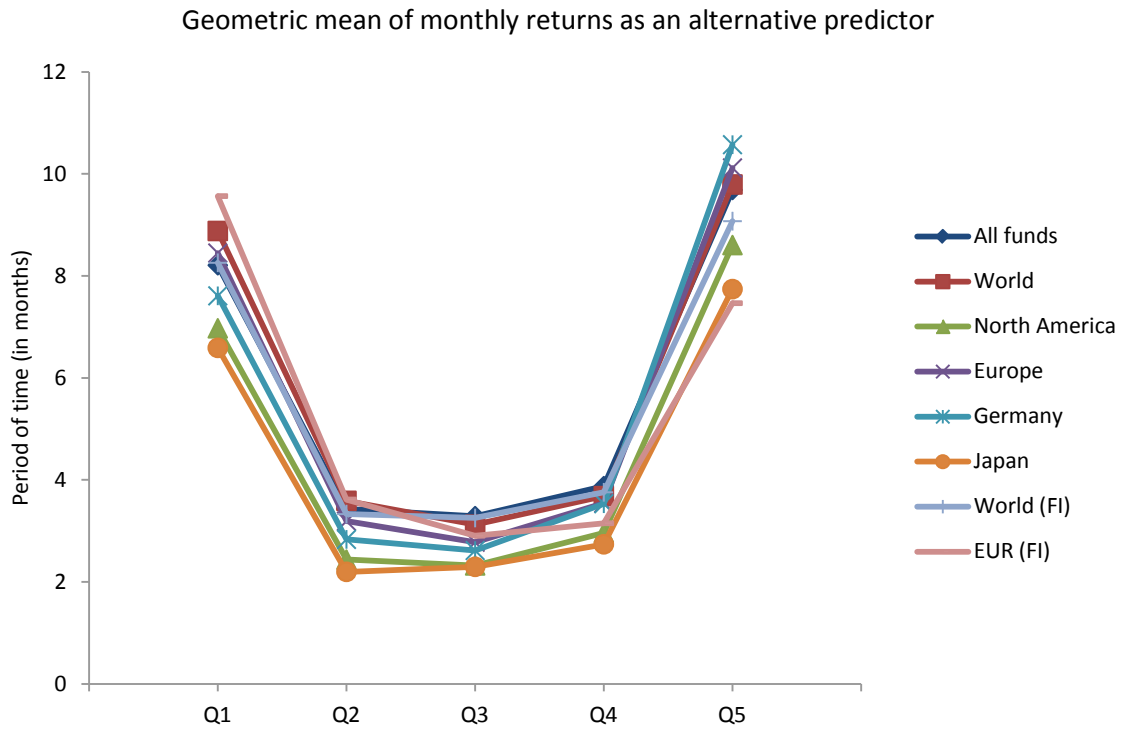
This table presents the results from the monthly dummy variable regressions of fund performance on the alternative predictor rating quintiles using the following performance measures: the index-adjusted return and the Jensen alpha. The performance is measured over the next 12 months after a given fund rating. Funds are rated monthly from December 2002 to December 2009. The post-rating performance is measured between December 2002 and December 2010. The category "All funds" includes all the 65 FondsNote fund rating categories. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. The *t*-statistics are adjusted by using the Newey-West procedure with a lag of 11 months. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

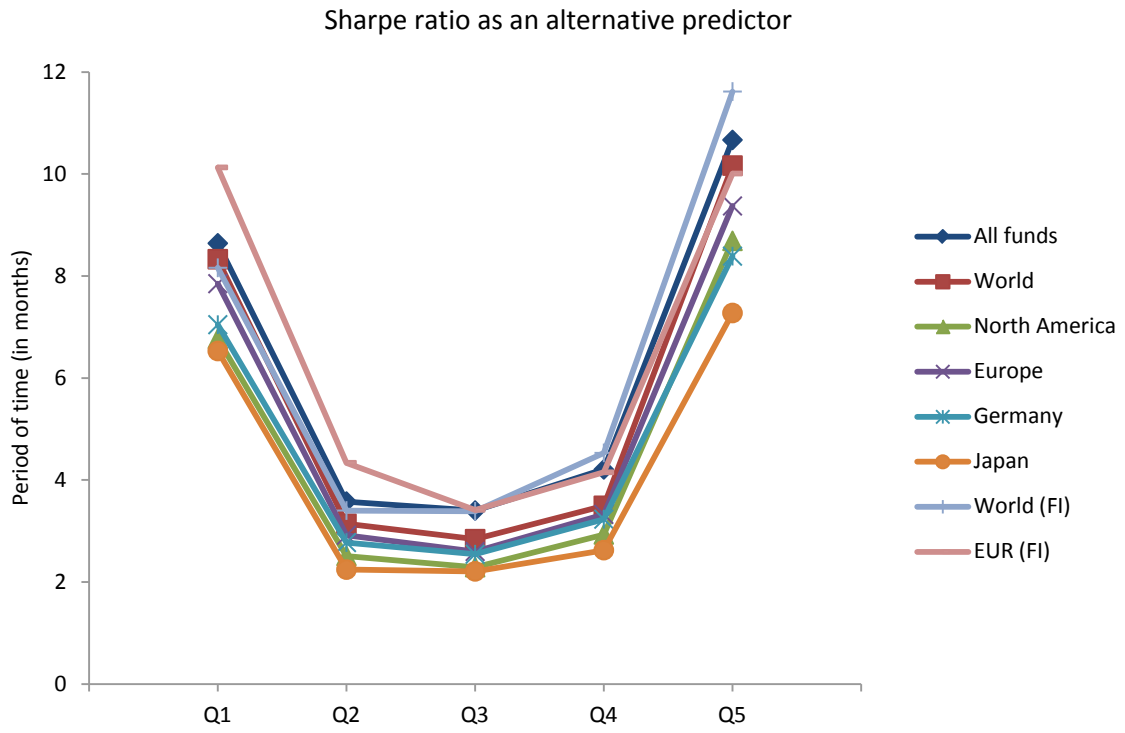
Table A5:
Performance Predictability of the FondsNote by Using the Index-Adjusted Return and the Jensen Alpha as Performance Metrics over the next 36 Months

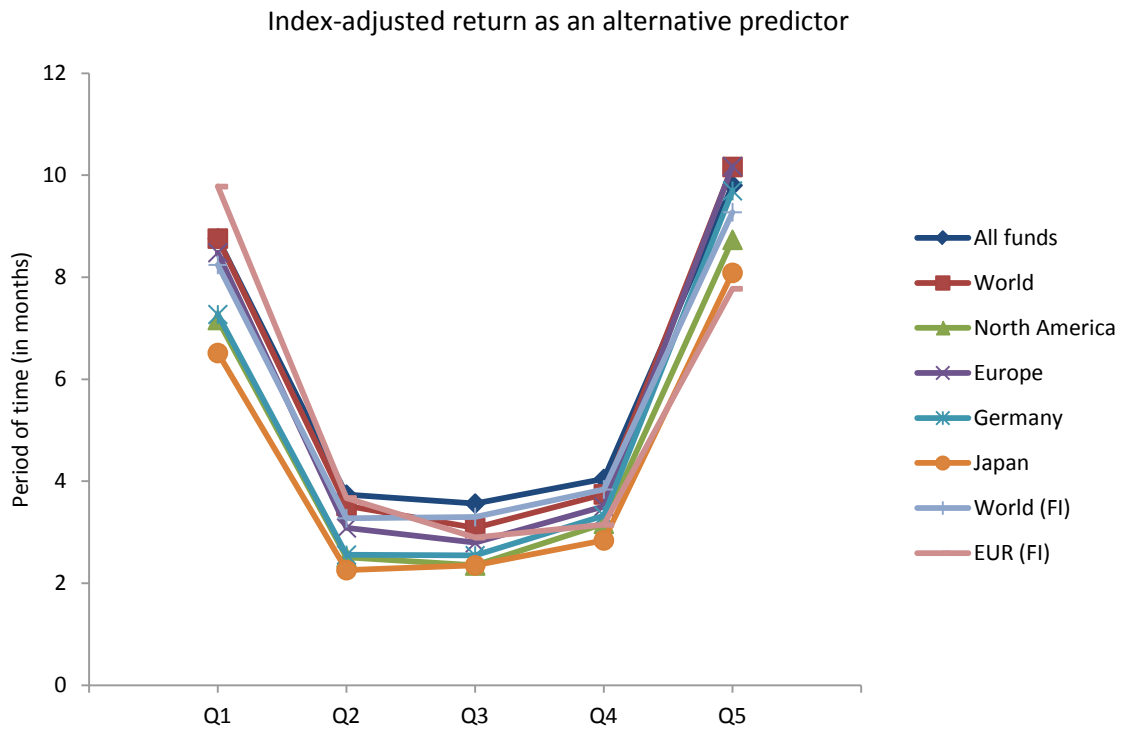
Fund sample	β_1	β_2	β_3	β_4	β_5	R^2	F-Stat.
Panel A: Index-adjusted return over next 36 months							
All funds	**0.0006*** **	0.0000	**0.0008***	**0.0022*** **	**0.0021*** **	0.29	**43.684*** **
	(3.624)	(0.046)	(1.876)	(3.342)	(4.509)		
World	**0.0027*** **	**-0.0025*** **	**-0.002*** **	**-0.002*** **	**-0.0011** **	0.12	**13.793*** **
	(6.650)	(-3.458)	(-5.094)	(-3.203)	(-2.171)		
North America	0.0001	-0.0004	**-0.0006***	-0.0003	0.0007	0.11	**12.424*** **
	(0.274)	(-1.075)	(-1.757)	(-0.619)	(0.594)		
Europe	**0.0005** **	-0.0003	0.0002	**0.0006** **	**0.0009*** **	0.14	**17.120*** **
	(2.507)	(-1.052)	(0.652)	(2.239)	(2.926)		
Germany	**0.0006** **	**-0.0008*** **	**-0.0006***	**-0.0014*** **	**-0.0024*** **	0.21	**27.576*** **
	(2.097)	(-2.800)	(-1.889)	(-2.932)	(-4.279)		
Japan	**0.0005*** **	0.0003	**0.0004** **	**0.0010*** **	**0.0011*** **	0.12	**14.967*** **
	(9.947)	(1.143)	(2.068)	(5.596)	(6.242)		
World (FI)	**0.0023*** **	**0.0004***	**0.0004*** **	**0.0008*** **	**0.0010*** **	0.18	**23.734*** **
	(21.154)	(1.884)	(2.768)	(3.884)	(5.297)		
EUR (FI)	**0.0034*** **	0.0000	0.0003	0.0004	0.0008	0.03	**3.535*** **
	(14.588)	(0.053)	(0.626)	(0.749)	(1.502)		
Panel B: Jensen's alpha over next 36 months							
All funds	**0.0006*** **	0.0002	**0.0010*** **	**0.0017*** **	**0.0018*** **	0.37	**62.531*** **
	(7.639)	(0.967)	(3.679)	(5.301)	(5.759)		
World	**0.0023*** **	**-0.0022*** **	**-0.0019*** **	**-0.0014***	-0.0009	0.10	**12.063*** **
	(3.391)	(-2.614)	(-2.591)	(-1.734)	(-1.082)		
North America	**-0.0003*** **	**-0.0004*** **	-0.0001	0.0002	**0.0009** **	0.12	**14.920*** **
	(-3.485)	(-2.959)	(-0.815)	(1.122)	(2.258)		
Europe	**-0.0003*** **	**0.0004** **	**0.0003***	**0.0011*** **	**0.0017*** **	0.24	**33.128*** **
	(-4.969)	(2.324)	(1.726)	(5.920)	(18.799)		
Germany	**0.0005***	**-0.0009*** **	**-0.0010** **	**-0.0017*** **	**-0.0028*** **	0.20	**25.850*** **
	(1.727)	(-2.688)	(-2.574)	(-3.889)	(-5.273)		
Japan	**0.0006*** **	**0.0006** **	**0.0006*** **	**0.0008*** **	**0.0014*** **	0.13	**16.264*** **
	(5.370)	(2.432)	(3.216)	(4.515)	(7.075)		
World (FI)	**0.0020*** **	**0.0003***	**0.0003** **	**0.0011*** **	**0.0015*** **	0.45	**86.098*** **
	(25.644)	(1.752)	(2.131)	(3.856)	(11.748)		
EUR (FI)	**0.0022*** **	0.0001	0.0001	**0.0007** **	**0.0012*** **	0.20	**26.314*** **
	(10.321)	(0.241)	(0.387)	(2.528)	(4.059)		

This table presents the results from the monthly dummy variable regressions of fund performance on the alternative predictor rating quintiles by using the following performance measures: the index-adjusted return and the Jensen alpha. The performance is measured over the next 36 months after a given fund rating. Funds are rated monthly from December 2002 to December 2009. The post-rating performance is measured between December 2002 and December 2010. The category "All funds" includes all the 65 FondsNote fund rating categories. Fama-MacBeth regression coefficients are the time-series average of the monthly cross-sectional regression coefficients. The *t*-statistics are adjusted by using the Newey-West procedure with a lag of 35 months. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

Figure A1: Rating Persistence of Alternative Predictors

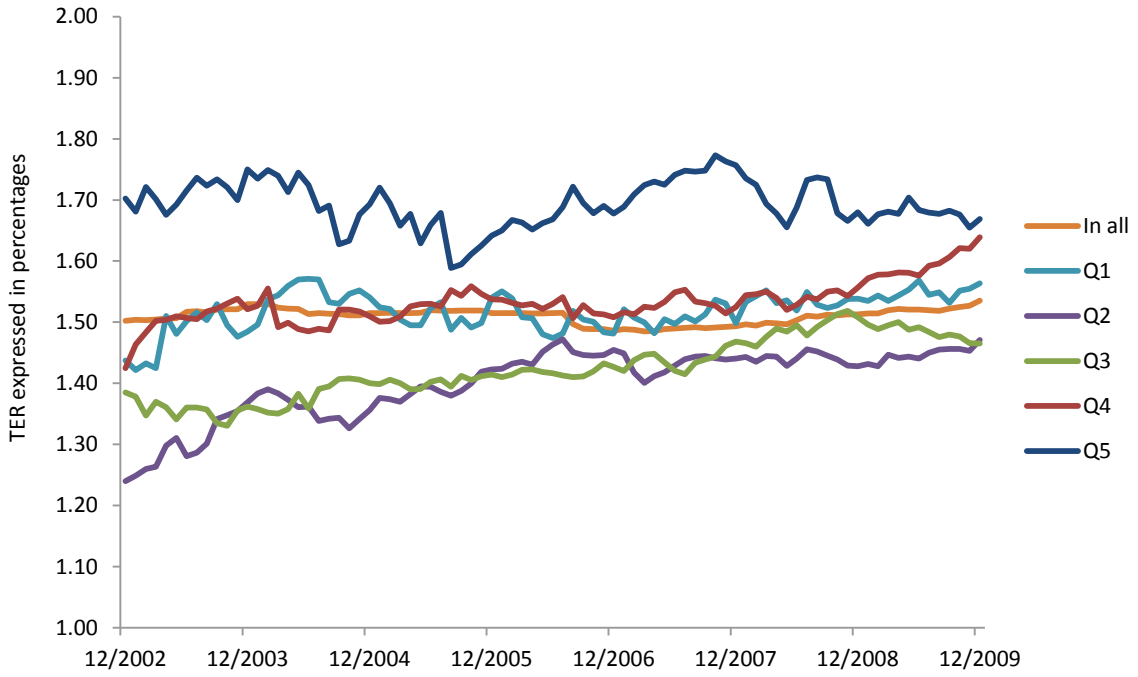






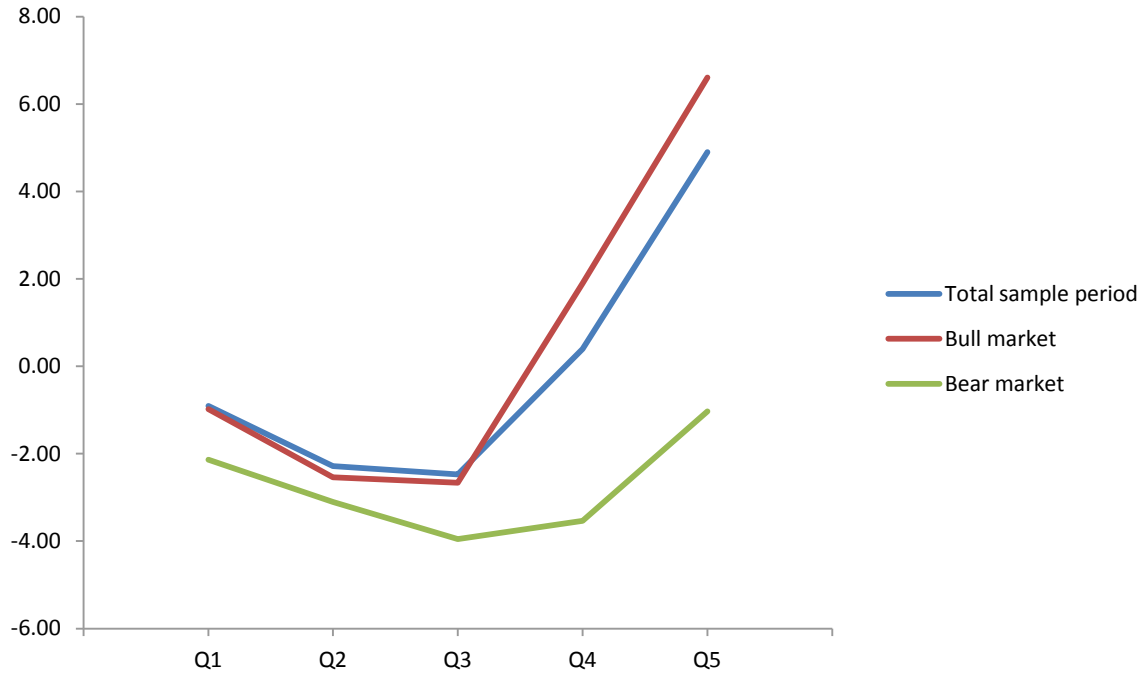
This figure presents the period of time (in months) till the rated funds change their ratings with a probability of 50%. The period of time is measured separately for the different rating classes. Fund ratings are based on the four-year in-sample performance of the geometric mean of monthly returns, the Sharpe ratio and the index-adjusted return. Quintile 1 (Q1) to quintile 5 (Q5) reflect funds that get the worst to the best fund rating at time 0. The category "All funds" includes the 65 FondsNote fund rating categories evaluated within the sample period from December 2002 to December 2009. The equity fund rating categories examined are: World, North America, Europe, Germany, and Japan. Fixed-income categories (FI) are: World and EUR.

Figure A2: Total Expense Ratios over Time



This table presents the relationship between fund rating classes and costs over time. Quintile 1 (Q1) to quintile 5 (Q5) reflect the funds that get the worst to the best fund ratings. Costs are defined as the total expense ratio of funds per year (arithmetic mean of fund rating classes) and are presented in percentages. The sample period comprises all the funds rated from December 2002 to December 2009. TERs are applied for the entire fund universe including the 65 different fund rating categories.

Figure A3: Fund Rating Classes and Fund Flows



This figure illustrates the relationship between fund rating classes of the FondsNote and the fund flows. Quintile 1 (Q1) to quintile 5 (Q5) reflect the funds that get the worst to the best fund ratings. Fund flows are presented in absolute terms (left axis, in millions of euros) on a monthly basis. The sample period comprises all the funds rated from December 2002 to December 2009. Moreover, two sub-periods are analyzed: the bull market (December 2002 to June 2007) and the bear market (December 2007 to February 2009). Fund flows are applied for the entire fund universe including the 65 different fund rating categories.

Selbständigkeitserklärung

Diese Dissertation basiert auf fünf Studien, die aus meiner Tätigkeit am Institut für Finanzierung resultieren. Die erste Studie „Physical and Synthetic Exchange-Traded Funds: The Good, the Bad, or the Ugly?“ ist in gemeinsamer Forschungsarbeit mit Frau Prof. Dr. Sigrid Müller und Herrn Dr. Stefan Schöne entstanden.

Ich habe die vorliegende Dissertation selbständig und ohne unzulässige Hilfe Dritter verfasst. Die statistischen Auswertungen wurden mit den Computerprogrammen SPSS und EViews durchgeführt, die Dissertation wurde in MS-Word erstellt. Außer diesen Computerprogrammen und der angeführten Literatur habe ich keine weiteren Hilfsmittel genutzt.

Ich habe mich bisher noch keinem anderen Doktorexamen unterzogen und habe die vorliegende Dissertation auch keiner anderen Fakultät oder einem ihrer Mitglieder vorgelegt.

Ich bezeuge durch meine Unterschrift, dass meine Angaben über die bei der Abfassung meiner Dissertation benutzten Hilfsmittel, über die mir zuteil gewordene Hilfe sowie über frühere Begutachtungen meiner Dissertation in jeder Hinsicht der Wahrheit entsprechen.

Christian Meinhardt