STOCKHOLM SCHOOL OF ECONOMICS
Department of Finance
Master Thesis in Finance
Spring 2009

Tutor: PhD Francesco Sangiorgi
Date: 2009-05-29, 13.15-15.00
Location: Room 750, SSE
Discussants: Filip Andersson, Niklas Westermark

The Effect of Morningstar Rating Changes on Mutual Fund Performance and Strategy

Christoffer Brenning[α] Viktor Fritzén[β]

Abstract

The purpose of this thesis is to examine if and how changes in Morningstar ratings affect the performance and investment strategy of mutual funds. This study uses a sample of 223 U.S. stock mutual funds that during the period January 2002 – February 2009 had at least one isolated Morningstar rating change over a two year period. The effect of Morningstar rating changes is analyzed as an event study. The benchmark for evaluating fund performance and investment strategy is the Carhart (1997) four factor model. It is found that there is a significant change in mutual fund total net assets following a Morningstar rating change, but the effects on performance and investment strategy are less pronounced. In the year following a rating change, there is an improvement in performance for upgraded funds and to a lesser extent a deterioration in performance for downgraded funds. This thesis finds that fund managers receiving a higher rating tend to reduce the risk of their portfolios in the year following a rating change while the effect on managers of downgraded funds is more ambiguous.

Acknowledgements: We would like to thank our tutor Francesco Sangiorgi for his guidance and input in writing this thesis. We would also like to thank Per-Olov Edlund and Cristian Huse for taking the time to discuss technical issues. Finally we would like to thank Paolo Sodini for providing us with access to our data source.

α 20304@student.hhs.se
β 20513@student.hhs.se

Table of Contents

1. Introduction 2

1.1 Background 2
1.2 Purpose of the Thesis 3
1.3 Relevance 3
1.4 Contribution to the Literature 3
1.5 Definitions and Abbreviations 4
1.6 Structure of the Thesis 4

2. Theoretical Background 5

2.1 The Efficient Market Hypothesis and the Presence of Fund Managers 5
2.2 The Impact of Fund Size on Performance 6
2.3 The Morningstar Effect 7
2.4 The Incentives of Fund Managers to Shift Risk 9

3. Hypotheses 11

4. Data Description 12

4.1 Sample Data 12
4.2 Sample Period 13
4.3 Variables 13
4.4 Sample Selection Methodology 15
4.5 Data Features 15

5. Methodology 16

5.1 Morningstar Rating Changes and TNA 16
5.2 Event Window and Estimation Period 16
5.3 Sorting of Cases 18
5.4 Regression Model Choice and Robustness 18
5.5 The Regression Model 19

6. Results 20

6.1 The Effect of Morningstar Rating Changes on Fund TNA 20
6.2 General Comments on Regression Results 21
6.3 First Sorting – UP and DOWN 22
6.4 Second Sorting – Rating Change Type 23
6.5 Third Sorting – TNA 24
6.6 Results in Relation to Hypotheses Two to Six 25
6.7 Limitations of the Results 26

7. Conclusions 27

8. Suggestions for Future Research 28

9. Reference List 29

9.1 Academic Sources 29
9.2 Non-academic Sources 32

10. Appendix 33

1. Introduction

1.1 Background

Morningstar is one of the leading financial information intermediaries offering Internet, software, and print-based products and services for its clients which include private investors, financial advisors and institutional investors (Morningstar, 2009). It provides data on 300 000 investment offerings ranging from stocks, hedge funds and other investment vehicles. But Morningstar is most famous for its system of star ratings of mutual funds with its five degree scale. To demonstrate superior performance, mutual fund companies frequently emphasize Morningstar ratings in the marketing of their funds, instead of their own return history, attesting to the popularity of Morningstar ratings among retail investors (Jones and Smythe, 2003). Its universal appeal has allowed Morningstar to expand globally and today the research company operates in 18 countries worldwide. Indeed, studies have shown that Morningstar's ratings have significant affects on investors in Finland (Knuutila et al., 2007), Denmark (Bechmann and Rangvid, 2007), and Australia (Gerrans, 2004).

Research has shown that investors care a great deal about what Morningstar rating mutual funds are awarded. In a study by Capon et al. (1996), the authors found that many mutual fund owning households consider published performance rankings as the most important source of information in their mutual fund investment decisions. Further, it has been found by Del Guercio and Tkac (2008) that the discrete change in Morningstar ratings cause investors to rebalance their fund portfolios, which has been dubbed the "Morningstar effect". The question of *how* influential Morningstar ratings are on the asset allocation decisions of investors and fund managers and *how* effective the Morningstar rating system is as a means for investors to earn superior returns has been the subject of much research, most notably by Blake and Morey (2000), Morey (2003), Gottesman and Morey (2006) and Del Guercio and Tkac (2008). Morey (2003) found that after receiving an initial 5 star rating, mutual funds experienced a significant deterioration in performance, coupled with changes in loadings on the Carhart (1997) four factors.[1] The author concludes that fund managers significantly alter the composition of their portfolio in response to an initial top rating, and that investors should be wary about using a 5 star rating as a predictor of future performance. Morey's study, however, is limited to funds receiving an initial 5 star rating, which is a special case. Further, Morningstar changed its rating system in 2002 (discussed in Section 2.3), and Morey's study is based on a dataset comprised of mutual funds between 1993 and 2001, prior to this change in methodology. This begs the question if Morey's findings are confined to the old Morningstar rating system and only to funds receiving an initial 5 star rating or if these effects are universal for rating

[1] See Section 1.5 Definitions and Abbreviations for a description of these factors.

changes in general also under the new rating system? If so, what factors drive this alteration in investment strategy and performance deterioration? These questions will be addressed in coming sections.

1.2 Purpose of the Thesis

In order to understand Morey's findings it is necessary to first understand why and how fund performance and investment strategy may be affected by changes in Morningstar ratings. Two research bodies in finance that we believe can potentially shed light on these issues are explored. The first is the research examining the impact of fund size on performance. The second deals with the incentives of fund managers to shift risk in response to past performance. This thesis attempts to link the above research bodies together by considering Morningstar rating changes within the framework of fund size and the incentive structure of fund managers. Specifically, the purpose of this thesis is to examine *if and how changes in Morningstar ratings affect the performance and investment strategy of mutual funds*.

1.3 Relevance

Given the popularity of the Morningstar rating system among both laymen and professional investors, there is strong reason why there should be an interest in this area. From a practical point of view this thesis contributes to the understanding of what investors should expect given changes in Morningstar ratings, which can be extrapolated for practical purposes to useful insights on appropriate courses of action following a rating change. From a theoretical point of view this thesis is relevant as it contributes to the literature on the interrelation amongst size and performance and strategy, and to the growing number of articles on Morningstar.

1.4 Contribution to the Literature

In the context of previous finance literature, this thesis distinguishes itself by examining how *changes* in Morningstar ratings affect performance and investment strategy of fund managers. Thereby, the direct casual link is studied, as opposed to examining differences between fund star groups in the cross section. To establish that the Morningstar effect as defined by Del Guercio and Tkac (2008) is still present under the new Morningstar ratings system we perform a simple test in order to determine if the implications of their findings are still relevant. More importantly, this thesis contributes to the current literature by extending the research by Morey (2003) to include all funds experiencing a Morningstar rating change, not just the segment of funds receiving an initial 5 star rating. The dataset in this study consists of mutual funds awarded Morningstar ratings based on the new rating system, which makes for interesting comparisons with studies performed with data on the old system. And to date, to the best of our knowledge, no study has been performed on the effect of changes in Morningstar ratings on performance and strategy under the new Morningstar rating system. Overall, the scope of this study is wider than the

one performed by Morey (2003) allowing for more general inferences about the effects of Morningstar rating changes on funds and fund managers.

1.5 Definitions and Abbreviations

When "performance" is referred to in this thesis, it is analogous to "alpha", which is a widely used performance metric. In regression models, alpha is measured by the intercept, which is the average excess return that is unaccounted for by the model. This is interpreted as a measure of the fund manager's level of "skill". Throughout the thesis, "performance" and the term "abnormal return" are used synonymously.

The regression model in this thesis is based on the Carhart (1997) four factor model. These factors are described in more detail below and in the data section. It should be noted that these factors are sometimes called "risk factors".[2] In this thesis, we will refer to the factors collectively as "the four factors".

The four factors consist of the Fama and French (1993) three factors; the market return in excess of the one month U.S. T-bill rate (RMRF), "Small minus Big" (SMB), "High minus Low" (HML), and the Jegadeesh and Titman (1993) momentum factor (MOM). The factors SMB, HML and MOM measure the differences in returns of portfolios of small stocks and large stocks, value stocks and growth stocks, and past winning stocks and past losing stocks respectively.

There are several different terms for what in this thesis will be referred to as "strategy" or "investment strategy". In performance attribution, "strategy" relates to loadings on factors e.g. the ones named above. "Strategy" is synonymous to the term "factor loadings", and both are used interchangeably in the study. In some other research this has also been referred to as "style", but that term will not be used.

The term "TNA" denotes total net assets and is used interchangeably with fund size.

The term "Morningstar effect" in this thesis refers to the direct effect following a Morningstar rating change on TNA as well as other secondary effects this might entail and other perhaps intangible effects. This definition is thus used in a somewhat wider sense than in previous research.

The term "fund" is used synonymously with "mutual fund".

1.6 Structure of the Thesis

The thesis is organized as follows; Section 2 provides an overview of previous research that we believe to be relevant, and lays the foundation for our theoretical framework from which we formulate our set of

[2] There is however an unsettled debate on whether some of these factors truly are proxies for aggregate sources of risk of concern to investors, anomalies resulting from market inefficiencies or the product of data mining. See Cochrane's (1999) "New facts in finance" for an overview of this subject.

hypotheses in Section 3. Section 4 describes the data used to test the hypotheses. The methodology in Section 5 outlines the selection and sorting procedures of the study and specifies the regression model used. The results of the study are presented in Section 6 and Section 7 concludes. In the final Section 8 suggestions for future research are proposed. The Appendix includes output from the regressions and other relevant data.

2. Theoretical Background

2.1 The Efficient Market Hypothesis and the Presence of Fund Managers

The ability of fund managers to consistently beat the market year after year has been the subject of much research in the empirical finance literature. Carhart (1997) and Fama and French (2008) show that there is some persistence in performance, but most of this persistence is explained by funds loading on the Carhart (1997) four factors. Once these factors are controlled for, the average fund generates zero or even negative abnormal returns net of fees, and most of the persistence in performance disappears over time.

Many have interpreted these findings as evidence supporting the Efficient Market Hypothesis. If markets are informationally efficient no single agent should consistently be able to beat the market on a risk-adjusted basis. Therefore, a fund manager that performs well one year is no more likely to perform well in the next. If this were the case, proponents of the Efficient Market Hypothesis argue that the observed persistence is due to luck, not skill on the part of the fund manager. In this framework, fund managers should on average underperform the market by an amount corresponding to the fees they charge for their services and the transaction costs incurred from their trading activities. But if this were true, no rational investor would choose to invest in actively managed funds. For investors to be indifferent between investing in a passive index fund and an actively managed fund, the risk-adjusted returns generated by the two, net of fees, need to be equal. Given this background, it is surprising that there exists a market for active fund managers at all. What is more, during the past 30 years, U.S. households have increased their exposure to the mutual fund industry. Today more than four in ten U.S. households have holdings in mutual funds. Further, it has been shown that the vast majority of U.S. retirement savings in defined contribution plans and individual retirement accounts are invested in actively managed funds (Investment Company Institute, 2008). This suggests that investors do have faith in the ability of active fund managers to add value.

2.2 The Impact of Fund Size on Performance

Do the above findings imply that the average investor is irrational in her asset allocation decisions? Not necessarily. Berk and Green (2004) propose that many commonly held views about fund managers and the fund management industry in general are, in fact myths. In their framework, the authors show that in competitive financial markets, where investors compete for the services of skilled fund managers, capital flows into actively managed funds as long as they are able to deliver higher expected returns than the investor could achieve for herself by investing in a passive market portfolio. However, because of diseconomies of scale in asset management, as funds grow larger, fund managers cannot effectively employ her skill at generating abnormal returns, resulting in deteriorating performance. In order for the market to clear, capital must flow into actively managed funds until the expected risk-adjusted return is the same as for a passively managed portfolio. The negative impact of size on performance means that competition among investors for skilled fund managers will drive out any ability for managers to earn abnormal returns. Several explanations for these diseconomies of scale have been proposed.

The link between size and performance has been studied extensively in the empirical finance literature for some time. Early studies by Grinblatt and Titman (1989) investigating this relationship found that gross of fees, smaller funds outperformed larger funds even after adjusting for risk. However, net of fees, they could not find the same result. Later studies by e.g. Chen et al. (2004) found that an increase in fund size is associated with poorer performance even after controlling for other fund characteristics such as fund turnover, age, expense ratio, etc. Further, they found that the adverse effect of size on performance is most pronounced among funds investing in small and illiquid stocks. Smaller funds are also significantly more likely than larger funds to invest in local stocks, and do so more successfully. The authors attributed these effects to costs associated with liquidity and price impact and organisational diseconomies arising from hierarchy costs.

Because the trading activities of large funds are more likely to impact prices than the trading activities of small funds, investments perceived as being profitable may not be carried out by large funds because of high transaction costs. Alternatively these trades might be carried out with longer execution time to avoid price impacts. Indeed, Indro et al. (1999) propose that transaction costs increase with fund size because trading of large volumes of stock drive up the bid-ask spread for stock prices significantly more than smaller volumes. This adverse affect may induce fund managers to defer trades or not carry them out at all. Delayed execution time of trades obviously impedes fund managers' ability to effectively time the market. This line of thought is very much aligned with Beckers and Vaughan (2001) who argue that "*every investment strategy eventually becomes self-defeating when too much money chases the same investment opportunity*" and go on to show that increased execution time of trades is negatively correlated

with fund performance. The above findings suggest that as funds grow in size it becomes more costly to deviate from the market index, hence funds that become very large eventually start to mimic a benchmark themselves.

Chen et al. (2004) argue that small funds to a greater extent are able to invest all their capital in their best ideas whereas large funds face greater liquidity constraints forcing them to invest in their not-so-good ideas and take larger positions in single stocks than is optimal, resulting in poorer risk-adjusted performance. This implies that as funds grow in size they need to generate more good investment ideas to retain their past performance. This seems plausible since funds that grow in size can afford to hire new staff to research a larger universe of investment opportunities. If funds respond to large capital influxes by hiring additional staff, one might expect that larger funds have a significantly larger number of holdings than smaller firms. However, this does not seem to be the case. Pollet and Wilson (2008) find that funds overwhelmingly increase their stake in existing holdings in response to growth in assets under management and diversify mainly to compensate for liquidity constraints. As an example they point out that in 2000, the average large fund held fewer than twice the number of stocks in its portfolio compared to a fund one hundredth its size.

However, it should be stressed that fund size can also provide benefits. Indeed, Indro et al. (1999) argue that growth in assets under management can initially provide cost advantages as mutual fund expenses such as costs to access data, research services, administrative and overhead expenses increase at a slower rate than fund size. Similarly, Latzko (1999) proposes that since many mutual fund expenses are fixed, economies of scale can arise from a fund's ability to distribute its fixed costs over a larger amount of capital. However, the author finds that these benefits are largely exhausted once a fund's assets under management reach $3.5 billion, implying that there may exist an upper limit in fund size for achieving economies of scale. The discussion of economies and diseconomies of scale in asset management suggests a trade-off between the two which begs the question of what the "right" amount of assets under management is. This question is addressed by Perold and Salomon (1991) who argue that the "right" amount of assets under management for a fund will primarily depend on the quality of its research, its transaction needs and the liquidity of the markets in which it trades. Further, the right amount of assets under management may vary significantly across different investment processes. The right amount of capital for a fund following one investment strategy may thus be suboptimal for another.

2.3 The Morningstar Effect

Let us assume, as do Berk and Green (2004), that investors cannot initially distinguish skilled from unskilled fund managers. Let us assume further that there are costs associated with gathering and

analyzing past performance data on fund managers. Further, investors view past performance as an indicator of managerial skill and, hence, ability to generate future abnormal returns. Given the vast universe of funds available to investors, this is undoubtedly a daunting task even for the professional investor. Instead of carrying out their own analyses, investors evaluate past performance based on other simpler methods, a popular one being Morningstar's mutual fund ratings. As Morningstar is *the* name within the mutual fund evaluation industry, its influence on investors' asset allocation decisions could be expected to be significant. Indeed, McGuigan (2006) shows that the stronger brand name an entity has, the less information investors choose to seek out for themselves. Further evidence of Morningstar's influence on investors includes Sirri and Tufano (1998) and Goetzmann and Peles (1997) who show that the top ranked Morningstar funds receive the largest cash inflows. Similarly, Keenan (2002) notes that 50.5% of all assets of U.S. equity funds in 2001 were claimed by 4 and 5 star rated funds. Further, seeing how mutual fund companies actively use Morningstar ratings in their marketing as a sales point, this indicates that investors care about a fund's Morningstar rating. As an example, Jones and Smythe (2003) found that out of 170 fund advertisements in Money magazine in 1999, 80 had some type of performance rating from an independent research institute. Out of these 80, 59 explicitly mention Morningstar by name.

If the assumptions of the Berk Green (2004) model hold, funds receiving a higher (lower) Morningstar rating should experience an inflow (outflow) of capital. This idea was tested on U.S. equity funds between 1996 and 1999 by Del Guercio and Tkac (2008) who find that changes in Morningstar ratings lead to abnormal fund flows up to as much as seven months following a rating change. Further, they find that it is the discrete change in the rating itself, and not changes in the variables underlying the rating, that causes the observed increase (decrease) in fund flows. The authors dub this causal phenomenon the "Morningstar effect". In this thesis, however, the concept is extended to include secondary effects, as was pointed out in the definitions section. This finding supports the idea proposed earlier that investors don't base their asset allocation decisions on independent analysis but instead update their expectation of fund managers' ability to generate abnormal returns using the Morningstar rating system.

At this point it is worth commenting on the change in Morningstar's rating system. Under the old rating system Morningstar calculated ratings for U.S. stock funds by comparing each funds' performance with all other domestic stock funds. The new system groups funds into 48 different categories that contain similar types of funds, as opposed to the old system which only had 4 broad groupings. An important consequence was that the old method of rating substantially favored growth stocks. For instance, during the IT bubble, many technology funds received top ratings because of increasing stock prices and low volatility at the time. The new system also contains a more theoretically appealing measure of risk. The

old system used performance in relation to T-bills to calculate risk, whereas the new system uses past volatility, with emphasis on downside risk, as its new risk metric. Del Guercio and Tkac (2008) argue that the new system better reflects managerial skill rather than which style has recently performed well. These fundamental changes in the rating system makes Morningstar ratings interesting to investigate, as there is not much research performed on it to date. For a complete discussion on the differences between the two systems see Gottesman and Morey (2006) and Del Guercio and Tkac (2008).

The Del Guercio and Tkac (2008) study was based on the old Morningstar rating system, but the authors "*expect that [their] main results would be qualitatively similar in a more recent period under the new rating methodology*". Although reproducing their results on current data using their methodology would be an interesting research topic, it is beyond the scope of this thesis. For the purposes of this thesis it suffices to establish that Morningstar rating changes do indeed lead to significant changes in fund size. This issue is addressed in later sections.

Under the old Morningstar rating system, Morey (2003) found that after receiving its initial 5 star rating, funds experienced a significant deterioration in performance the following three years. Interestingly, Del Guercio and Tkac (2008) found that that the "Morningstar effect" was most pronounced among funds receiving a 5 star rating, i.e. funds that received a 5 star rating had the largest abnormal inflows of new capital of all funds. These results are consistent with the predictions of the Berk and Green (2004) model that past winning funds experience an influx of new capital when investors flock to the most skilled fund managers to chase abnormal returns, resulting in deteriorating performance. Further, Morey (2003) found that after receiving its initial 5 star rating, fund managers significantly increased their loading on HML in order to retain their rating. At the same time, fund managers decreased their loading on MOM and increased their loading on the market excess return. However, in a later study Gottesman and Morey (2006) found that the new Morningstar system has power in predicting future fund performance.

2.4 The Incentives of Fund Managers to Shift Risk

Studies in the behavioral finance field have tried to link past performance to risk-taking behavior of fund managers. Several theories have been proposed to explain this relationship. One theory proposes that investors supply capital to fund managers, who compete with each other for new capital inflows. In this sense, the mutual fund industry is very much like a "tournament". Further, investors supply fund managers that have recently generated high abnormal returns with new capital but do not withdraw capital from fund managers who have recently performed poorly to the same extent. This gives rise to a convex relation between past performance and fund flows. Indeed, in a study by Sirri and Tufano (1998), the authors examine the relationship between past performance and fund flows using a dataset of U.S. equity

funds between 1971 and 1990. The authors conclude that investors disproportionately flock to winning funds but fail to flee loosing funds at the same rate. Assuming that a fund manager's compensation scheme is an increasing function of the size of the fund she manages, which is commonplace, this creates an incentive for the fund manager to increase (decrease) risk in response to poor (good) past performance. Fund managers that perform well in one period are more likely to "play it safe" and reduce the riskiness of their portfolio in the next period to lock-in their ranking relative their peers. By contrast, fund managers that perform poorly in one period bear more upside than downside risk and are thus more likely to "gamble" and increase the riskiness of their portfolio in an attempt to catch up to their peers. Indeed, the findings of Brown, Harlow and Starks (1996) lend support to this theory. However, in a more recent study, Ammann and Verhofen (2007) found that fund managers that had recently performed well increased the volatility of their portfolios and increased the loadings on SMB, HML and MOM to a greater extent than poorly performing fund managers. These instead increased their tracking error, i.e. the deviation of the funds return from their benchmark.

Can the findings of Ammann and Verhofen (2007) be reconciled with theory? Kale et al. (2008) develop another multi-period model in which fund managers are employed by fund companies and can be fired if performance is bad. Investors supply capital to fund managers they perceive as being skilled, and thus most likely to generate future abnormal returns, based on past performance. As in the Berk and Green (2004) framework there are diseconomies of scale in asset management. In the bargaining game between fund company and fund manager it is assumed that the fund manager's payoff is an increasing *convex* function of her perceived skill. Fund managers that have recently generated high abnormal return relative their peers are unlikely to be fired and thus choose to increase the riskiness of their portfolios to benefit from the convexity in their payoffs. By contrast, fund managers that have performed poorly are likely to be fired and hence increase the risk of their portfolios thereby increasing the probability of their perceived skill improving and not being fired. Average performing fund managers will choose lower risk levels out of fear of being fired compared to the best and worst performing fund managers. This interplay between the fund manager's compensation scheme and career concerns leads to a U-shaped relation between past performance and relative risk-taking of the fund manager. Using a dataset of U.S. mutual funds from 1996 to 2002, Kale et al. (2008) find a significant U-shaped relation in the data lending support to the idea that career concerns are a driving factor in risk-shifting by fund managers. These findings are also consistent with those of Morey (2003) that after receiving an initial 5 star rating, fund managers significantly increase the riskiness of their portfolios, as measured by both sigma and beta.

3. Hypotheses

From the theoretical discussion in the previous section we are now ready to formulate the hypotheses that this study aims to test. In light of the Berk and Green (2004) framework that investors compete for the services of the most skilled fund managers and the findings of Del Guercio and Tkac (2008) that investors flock to upgraded funds and flee downgraded funds under the old Morningstar system, we expect that the TNA of funds experiencing a rating upgrade (downgrade) will increase (decrease).

Hypothesis 1: *Funds receiving a rating upgrade (downgrade) will see their TNA increase (decrease)*

Assuming that there are diseconomies of scale in asset management, an increase (decrease) in TNA will cause fund managers' performance to deteriorate (improve). Several explanations for these diseconomies have been proposed; among them are large funds' inability to invest all their capital in their best ideas, costs associated with liquidity and price impact, and organizational diseconomies. Given that Hypothesis 1 holds, we expect that upgraded (downgraded) funds will experience a performance deterioration (improvement). In this light, we formulate our second hypothesis.

Hypothesis 2: *Abnormal returns decrease (increase) for funds receiving a higher (lower) rating*

Previous studies have shown that the "Morningstar effect" is more pronounced for some rating changes compared to others. This suggests that in the presence of diseconomies of scale, some funds' performance, depending on the rating they held before and after their rating change, i.e. their rating change type, will be more affected by a rating change than others. The discussion on fund size and performance showed that this relationship is multidimensional. The asset management industry seems to be characterized by both economies and diseconomies of scale suggesting that the fund size – performance relationship may be non-linear. For example, a small fund may be able to double its size many times over before it begins to feel constrained by its size in terms of carrying out desired investments. However, once these constraints begin to set in they become increasingly binding making it more difficult for fund managers to effectively employ her skill. But once a fund reaches a certain size, additional growth only marginally impedes performance. The largest funds are more likely to be constrained to the point that they to a larger extent choose to follow a passive investment strategy, tracking their benchmark compared to smaller actively managed funds. We therefore suspect that the effects of rating changes on performance may differ across funds of different size. This leads to our third hypothesis.

Hypothesis 3: *The effect of Morningstar rating changes on fund performance varies across cases sorted on rating change type and on fund size*

Several theories were proposed earlier on how fund managers respond to past performance. In light of the theoretical discussion on fund managers' inventive to risk-shift, we believe that changes in Morningstar ratings may induce fund managers to increase the riskiness of their portfolios in response to a rating downgrade by increasing loadings on the four factors.

Hypothesis 4: *Loadings on the four factors increase for downgraded funds*

For funds that experience a rating upgrade, fund managers are induced to change the riskiness of their portfolio. As theory provides arguments for both cases, this change can be manifested as either an increase or a decrease in the loadings on the four factors.

Hypothesis 5: *Loadings on the four factors change for upgraded funds*

Given that managers of large funds face greater costs associated with liquidity and price impact, they incur larger costs for deviating from the market benchmark and therefore load less on strategies other than the market excess return. Therefore, it is expected that they are likely to be less responsive to rating changes compared to smaller funds.

Hypothesis 6: *The effect on strategy of rating changes is more pronounced among small and intermediate sized funds than among larger funds*

These hypotheses are aimed at fulfilling the purpose of the thesis; to get an understanding of if and how Morningstar ratings affect the performance and investment strategy of mutual funds.

4. Data Description

4.1 Sample Data

The dataset in the study is comprised of U.S. equity funds denominated in U.S. dollars, domiciled in the U.S. Morningstar has four broad categories of mutual funds. The sample in this study consists only of funds within Morningstar's *Broad Asset Category* called *U.S. Stock*. The criterion for this group is that at

least 75 percent of the assets are invested in equities and that at least 75 percent of the value of the equities consists of investments in U.S. companies. These are both calculated as three year averages (Morningstar, 2007). The sample of funds in this study contains both actively and passively managed funds, because both are part of the investment universe for investors. The proportion of wealth in actively managed funds in the U.S. market was during the sample period approximately between 83 and 89 percent (Seeking Alpha, 2008). The *U.S. Stock* category is chosen for several reasons. Most importantly, it is the same category of funds used by Del Guercio and Tkac (2008) in their study of abnormal flows following Morningstar rating changes. It is also the fund category used in the study by Morey (2003). The reason for using a similar fund category as the studies above is for increased comparability with them. Other positive aspects of the category are that it is large in the number of funds and in terms of market value. The U.S. equity category is the most significant fund category for U.S. investors and is well diversified across industries.

4.2 Sample Period

The time period in the sample is from January 2002 to February 2009. In 2002 Morningstar changed its rating system, and the chosen time period is the longest possible to date with the new rating system. The observations are on a monthly basis. The choice of monthly observations is natural as ratings for funds are updated on a monthly basis. This gives a sample of 86 fund months. This includes a four and a half year long bull run from 2003 to mid 2007 and the crash from mid 2007 to the beginning of 2009. The sample thus includes periods of both high and low market volatility. We believe that it is desirable that the sample period is diverse making it more representative for a whole business cycle. This also makes our results more general and valid in the future as well. However, it can be debated if the past seven years really have been representative of a "normal" business cycle. We examined whether the risk premia on the Carhart (1997) four factors in our sample period differed substantially from historical levels. The average risk premia for the sample period were positive for all factors except the excess market return and were comparable to historical averages. The Sharpe ratios for the factors during our sample period were also comparable to historical values except for the HML factor, which was significantly higher in our sample period compared to historical averages due to very low volatility. Overall, however, it does not seem like our sample period is unrepresentative in relation to historical periods in terms of risk premia on the four factors.

4.3 Variables

The variables used in this study for the individual funds are net returns, Morningstar ratings and estimated TNA divided by the market value of the S&P 500, which makes the size measure normalized for comparisons between time periods. These have been obtained from Morningstar's online data program

Morningstar Direct[3], except the time series of the market value of the S&P 500, which was collected from Datastream[4] and are described in more detail below.

Morningstar's calculation of net return is determined each month by taking the change in monthly net asset value (NAV), reinvesting all income and capital-gains distributions during that month, and dividing by the starting NAV. Returns are not adjusted for sales charges (such as front-end loads, deferred loads and redemption fees) but does account for management, administrative, 12b-1 fees and other costs taken out of fund assets. Aside from the loads that investors may face, the return metric used by Morningstar accurately reflects the return that investors receive.

The Morningstar rating system is a relative system that has a fixed proportion of funds within each star rating category.[5] Within each category, the top 10 percent of funds, according to Morningstar's risk-adjusted performance measure, receive 5 stars. The 4 star group consists of the following 22.5 percent best performers. The 3 star group consists of the 35 percent average performing funds. The distribution is symmetrical so the 2 star group has 22.5 percent and the 1 star group 10 percent (Morningstar, 2008). The Morningstar rating is a risk-adjusted measure, which takes into consideration a utility function of an average risk-averse investor. Ratings are published at the beginning of every month. The most commonly referred to rating is the overall rating. This consists of a weighting of three different ratings, which are a three year, five year and ten year rating, for the funds that have existed that long. A fund that has existed a shorter period of time is given an overall rating based on the time period it has existed. Funds that are less than three years old are not rated (Morningstar, 2008). According to Blake and Morey (2000) an issue that arises when comparing funds on overall rating is that a fund that has only existed for three years is much more sensitive to the performance of the overall market during those three years. If those three years were a bull market period with low volatility, the fund will have a higher rating compared to a fund that is very similar, but has been around for ten years. A full length discussion on the age bias in Morningstar ratings can be found in Blake and Morey (2000) and Adkisson and Fraser (2003).

Estimated fund size is defined as the total amount of money managed as a standalone portfolio across share classes. We use estimated fund size because many funds only report TNA quarterly, and this is a monthly variable. For simplicity the term TNA is used throughout the thesis for the term estimated fund size. The use of TNA in the thesis is to be able to sort funds on a relative basis, making the need for an exact value between reported periods less important.

[3] Morningstar Direct Version 3.6.1
[4] Datastream AdvanceVersion 4.0
[5] To clarify on the terminology, Morningstar uses the term Broad Asset Category to define its largest four groups. In the subgroups the term category is also used and there refers to groups such as Small-cap, Value, Emerging Market Bonds etc. In this paper category will refer to this latter sub grouping, unless otherwise specified.

The reason for dividing TNA by the market value of the S&P 500 is to increase comparability between funds across time periods. The S&P 500 was chosen primarily because it is one of the most commonly used benchmarks covering almost 75 percent of U.S. listed equity value and is a good proxy for the U.S. stock market (Standard & Poors, 2009). For the regressions in the study that are sorted on TNA at the time of the event, TNA is divided by the market value of the S&P 500, to control for changes in the overall market value, so that funds that are large in June 2007 are large relative the overall market and comparable to funds that are relatively large in February 2008, which would otherwise be small in size compared to June 2007.

The factors used in our regression model are the Carhart (1997) four factors; the monthly returns of the market portfolio in excess of the one month U.S. T-bill rate, SMB, HML and MOM portfolios which have all been collected from Kenneth French's homepage (French, 2009).

4.4 Sample Selection Methodology

The sample obtained from Morningstar Direct includes all currently active mutual funds as well as mutual funds that have been liquidated, so that the sample is free from survivorship bias. When the data was collected in mid March 2009 the number of unique funds Morningstar Direct had on storage that were classified as *U.S. Stock*, denominated in dollars and that were U.S. domiciled from January 2002 to February 2009, totaled 4431. Of these funds, there were 941 that either had missing return data or missing estimated fund size data, and were therefore dropped from the sample. Of the remaining funds, 223 had at least one rating change that was preceded by nineteen months of the same rating and held their new rating for at least twelve months. 4 funds had two rating changes during the period that satisfied the criterion. The total amount of events in the sample therefore is 227. 6 rating changes that were adequate in terms of the period pre and post the rating change but had rating changes that were more than 1 star. These funds were excluded as they are special cases and do not constitute a group large enough to use with any statistical power.

4.5 Data Features

In the dataset of 992 funds which had all the necessary data to be included in the regression, and had held a rating for at least one year (where 227 of these funds satisfied the event period criterion and are used in the regression), the average number of rating changes was 1.5 times per year per fund, totaling 9 518 changes for the 992 funds over 85 months. The median fund changed ratings 1.3 times per year. The fund that changed ratings the most frequently changed ratings 27 times during the 86 month period, which is almost once every three months.

For the interpretation of the results to be robust, the statistical properties of the sample have been examined. Simple and partial correlations between regressors and variance inflation factors show quite clearly that the data does not suffer to any great extent from multicollinearity. A test for autocorrelation[6] was performed where the null hypothesis of no autocorrelation could not be rejected in any of the groups. A likelihood ratio test showed that the sample was not homoskedastic, so this has to be taking into consideration when estimating our regression model. These results are presented in the Appendix in Tables 6 – 9.

5. Methodology

5.1 Morningstar Rating Changes and TNA

To determine if there are significant changes in TNA following Morningstar rating changes, as in the data sample used by Del Guercio and Tkac (2008), we compare the average TNA as a fraction of the market value of the S&P 500 pre and post a rating change. We apply a double sided 2 sample t-test that compares the pre and post averages assuming unequal variances as volatility can differ substantially over a 24 month period. By normalizing fund size we are to some extent able to control for the increase or decrease in size of the overall market. The change in TNA that is not accounted for by market movement has not directly been controlled for. This can potentially be a concern for the test. The mutual fund industry has seen its capital under management double from 2002 to 2007 (Investment Company Institute, 2008), where a significant part of that growth is due to new capital going into mutual funds, and not just because the value of existing holdings has increased. If this is a problem, it is then to be expected that this manifests itself in that upward rating changes will be highly significant and of large magnitude, while downward rating changes are insignificant or of small magnitude.

5.2 Event Window and Estimation Period

Given the event-type nature of the study, the first order of business is to define the event, event window and estimation period. The term "event" refers to a rating change that satisfies the criterion defined in Section 4.4. One must consider that Morningstar updates its ratings on a monthly basis, which means that a single fund may experience both up and downgradings in its rating within a fairly short time interval. Indeed, funds on average experienced 1.5 rating changes per year during our sample period. Studies discussed in previous sections have shown that changes in Morningstar ratings give rise to very different (potentially off-setting) effects. In order to study these effects in isolation, the event periods of individual

[6] Panel data test for autocorrelation developed in Wooldridge (2002) p. 282-283.

rating changes cannot be overlapping. The size of the event window chosen in this study is influenced by the results presented in the Del Guercio and Tkac (2008) article on fund flows, as well as a trade-off between efficiency in the results and sample size.

According to Del Guercio and Tkac (2008), abnormal flows to and from mutual funds are observable up to seven months after a rating change. Therefore, seven months is the minimum viable length of the event window. However, the restriction of a constant rating for seven months results in many funds being sorted away, since many funds have multiple rating changes within a seven month interval. On the other hand, in order to obtain efficient estimates of the coefficients, seven months is quite short, especially considering the number of variables in our regression model. In the compromise between efficiency and representativeness, a good middle ground is found where the marginal increase in efficiency is more than offset by the decrease in sample size. With the data used in this study, 12 months seems to be where this tradeoff is at its best. It might not be appropriate to use an even longer event window to increase efficiency, as it, in addition to decreasing the sample size, also increases the risk of other exogenous factors influencing the data being examined. The estimation period in the study is also twelve months. The choice of an estimation period that is the same length as the event window is, aside from the symmetry, somewhat arbitrary, but is also in the end a compromise between efficiency and sample size.

To ensure that the estimation periods of twelve months for the funds is free from abnormal flows, the seven months preceding the estimation period do not have any ratings changes, as Del Guercio and Tkac (2008) have shown that abnormal flows following a rating change can persist up to seven months following a rating change. Therefore, for each event in the study, there are no rating changes the preceding 19 months and no rating changes the 12 months following the rating change, which aggregates to a period of 31 months for each event period. Employing this method, the first event in the data is in September 2003 and the last event in the data is in February 2008.

At this point it should be stressed that rating upgrades (downgrades) do not mean that performance has improved (deteriorated) at the time of the rating change. The overall rating is a weighted average of performance in the last three, five, and ten years, i.e. the historical performance of a fund. This means that once there has been a "real" improvement in performance, e.g. because the research of a fund starts to bear fruit or a fund hires a new more skilled manager, it will take some time before this performance improvement is translated into a rating upgrade. Put differently, rating upgrades (downgrades) lag performance improvement (deterioration). Once a rating change does occur, funds have already been higher or lower performers for some time. Therefore, there is no rule dictating that in the year before and after a rating change an upgraded (downgraded) fund must experience an improvement (deterioration) in

performance. This is because the improvement (deterioration) has likely already taken place prior to our event period. Because of the lag between a real performance change and a Morningstar rating change, we do not believe that our study is subjected to a look-ahead bias to any greater extent.

5.3 Sorting of Cases

To test our hypotheses, the 227 unique events have been sorted three times into different cases. The first sorting of the data is based on whether a fund has had an upward rating change or a downward rating change. These cases have been named UP and DOWN respectively. Given the findings by Del Guercio and Tkac (2008) that different rating changes are associated with differing degrees of abnormal flows, the second sorting is based on the change in the star group, referred to as rating change type. The first case is for funds that have had a 1 star rating and have been upgraded to a 2 star rating, which has been named 12. The second case consists of funds that have been upgraded from a 2 star rating to a 3 star rating and has been named 23, and the rest of the cases follow this method. In total for the rating change type sorting there are eight cases; four upgraded cases and four downgraded cases. The third sorting is based on a two step method. First, just as in the first sorting, the funds have been divided into those that have had an upgrade in rating and those that have had a downgrade in rating. Within these two groups, funds have been sorted into quintiles based on their market value in relation to the S&P 500 at the time of the event, UP1Q (DOWN1Q) being the upgraded (downgraded) funds with the lowest average market value and UP5Q (DOWN5Q) the largest. This third sorting methodology generated ten cases. All the cases and number of events for each case can be seen in Table 5.1 below.

Table 5.1 Cases and sample size

Case	Total	UP	DOWN	12	23	34	45	54	43	32	21
# events	227	84	143	10	32	33	9	19	63	50	11
# funds months	5448	2016	3432	240	768	792	216	456	1512	1200	264

Case	UP5Q	UP4Q	UP3Q	UP2Q	UP1Q	DOWN5Q	DOWN4Q	DOWN3Q	DOWN2Q	DOWN1Q
# events	16	17	17	17	16	28	29	29	29	28
# funds months	384	408	408	408	384	672	696	696	696	672

5.4 Regression Model Choice and Robustness

The choice of model was guided by econometrical tests.[7] Breusch-Pagan and Haussman tests were performed for all of the cases in order to establish the most appropriate regression model to choose. These tests indicated that the most appropriate model was a Random Effects model, which also suits the

[7] For a more in depth discussion of the econometrics behind the regression, and model choice tests the reader is recommended to look into an econometrics text book. The authors in this thesis referred to Gujarati (2003) and Hsiao (2003) for initial consultations.

assumptions about different performance and managerial skills best. To further establish which model is the most appropriate for the dataset, the regressions were performed using four different models; a Generalized Least Square model, Random Effects model and Fixed Effects model (all being panel data models), and individual time series regressions for each event which are then averaged as used by Carhart (1997), among others. For the panel data models the results turned out to be very similar in terms of coefficients and test statistics across cases and models, suggesting that the choice between different panel data models is not crucial for the results. The signs were in almost every case the same across the three models, and the significance levels only varied marginally. The distribution of the error terms in the panel data regressions (regardless of which model was used) cannot be seen as normal due to fat tails and a few extreme outliers. However, the distributions are highly symmetrical, which still makes statistical inferences possible. The time series regressions yielded coefficients similar to those in the panel data models. Especially those coefficients that were significant in the time series regressions differed only slightly from the corresponding coefficient in the panel data models. The drawback with the time series regression model, however, is that the number of observations relative the number of endogenous variables is not large enough to produce coefficients with low enough variances for statistical inferences. The time series regression approach is theoretically appealing though in that it is not affected by a potential heterogeneity bias in the coefficients that may arise when data is pooled. However, the Random Effects model, due to the varying intercept, is able to mitigate some of the heterogeneity bias, and still produce coefficients with better significance levels.

5.5 The Regression Model

Recent performance attribution studies have used the Fama and French (1993) three factor model and later the extended Carhart (1997) four factor model which also includes the return from a momentum zero-net-investment strategy, which is the one used in this study. These trading strategies are by no means exhaustive but are widely recognized and well documented in the empirical finance literature. Because the purpose of this study is to examine whether changes in Morningstar ratings are associated with changes in funds' performance and investment strategy, dummy variables are included in the regressions to determine if there are any significant differences in the alpha and loadings on the different investment strategies before and after a rating change. The regression model used in the study is specified below

$$\begin{aligned} R_{it}^{e} = {} & \alpha_i + b_1 RMRF_t + s_1 SMB_t + h_1 HML_t + m_1 MOM_t + d_1 D_{it} + \\ & + b_2 D_{it} RMRF_t + s_2 D_{it} SMB_t + h_2 D_{it} HML_t + m_2 D_{it} MOM_t + e_{it} \end{aligned}$$

where R_{it}^{e} is the return of fund i in month t in excess of the one month U.S. T-bill rate, RMRF is the return on the market portfolio in excess of the one month U.S. T-bill rate, SMB (small minus big), HML (high minus low), MOM (momentum) are the value-weighted returns of zero-net-invest, factor mimicking portfolios for size, book-to-market equity and one year momentum in stock returns. For a complete description of the construction of the factor mimicking portfolios see Kenneth French's website (French, 2009). The dummy variable D takes the value 0 in the months prior to a rating change (estimation period) and 1 in the months following the rating change (event window). The intercept is the average excess return that is unaccounted for by the factors, which can be interpreted as a measure of the fund manager's level of "skill". The slope coefficients of each factor measure funds' average loadings on that particular factor. The slope coefficients of the factor dummies measure the average change in loadings on each factor in the year following a rating change.

As mentioned in Section 4.5, the sample data does not suffer from autocorrelation for any of the cases, but almost all of the cases are heteroskedastic. We have therefore controlled for heteroskedasticity in the regression using robust variances. This leaves the coefficients less significant, but more reliable.

6. Results

6.1 The Effect of Morningstar Rating Changes on Fund TNA

The results of the t-test for differences in average TNA relative the S&P 500 during the 12 month pre and post a rating change are presented in Table 1 in the Appendix. Table 2 also shows the average TNA in the year before and after a rating change for all the cases. However, these figures can be misleading as they do not consider the impact of change in market value during different periods. For upgraded funds, TNA as a fraction of the market capitalization of the S&P 500 increases significantly and the opposite is true for downgraded funds. This is consistent with the findings of Del Guercio and Tkac (2008). Worth noting is that the magnitude of the increase (decrease) in TNA for upgraded (downgraded) funds differ substantially between the two cases as both the average and median t-statistic show. The decrease in TNA for downgraded funds is much more pronounced than the increase in TNA for upgraded funds. These findings differ from those of Sirri and Tufano (1998) who found that high performing funds reap exceptionally high rewards in terms of capital influxes but investors fail to withdraw capital from poorly performing funds to the same extent. However, in the upgraded fund category, funds in case 12 reduce their TNA both in absolute terms and in proportion to the S&P 500 suggesting that investors flee from these funds, which is contrary to expectations. It is hard to find a plausible explanation for this based in theory, although it should be stressed that the sample size of this case is rather small, consisting of only

ten events. We suspected that due to the small sample size one explanation to this finding could be that these rating changes took place during 2007 and 2008 when stock markets crashed. If these funds were high risk, a stock market crash coupled with the "flight to quality" that one might expect during bear markets would have a particularly adverse affect in terms of driving down the value of the assets held by those funds. However, upon closer inspection it turned out that most of the rating changed for this case took place prior to the crash, so this explanation fails. Case 12 could explain why the magnitude of the increase in size for upgraded funds is less pronounced compared to decrease in size for downgraded funds. In case 54 the average t-statistic has the wrong sign, whereas the median has the correct one. Neither is significant though. This is in line with what Del Guercio and Tkac (2008) found, that investors do not punish funds that are downgraded from 5 to 4 stars which they deem as a "non-event". For some of the cases, the sample sizes are not sufficiently large to draw statistically significant conclusions. However, in the large samples, the results are quite robust.

That the magnitude for downgraded funds was larger than for upgraded funds shows that the concern raised in the methodology section that the growth of the mutual fund industry would lead to poor t-statistics for the downgraded cases was not materialized. Overall, the t-tests indicate that there is a significant impact on TNA following changes in Morningstar ratings, not beyond any doubt, but to the extent that the intuition from Del Guercio and Tkac (2008) should be applicable to the dataset in this study. This simple test does not consider many other factors that might be affecting the fund size. Del Guercio and Tkac (2008) perform a much more rigorous study to demonstrate and isolate the Morningstar effect on flows. The purpose of these tests is merely to show that the dataset used in this study is impacted by Morningstar ratings changes, which is crucial for the rest of the tests. Overall, these findings support Hypothesis 1 that Morningstar rating upgrades (downgrades) do indeed lead to a significant increase (decrease) in funds' TNA.

Key Finding

- **Rating upgrades (downgrades) lead to an increase (decrease) in fund total net assets as a fraction of S&P 500 market capitalization.**

6.2 General Comments on Regression Results

The results for the three sortings are discussed in the three following sections below. For each section, the first paragraph deals with abnormal returns, and the second paragraph deals with strategy. Generally, the four factor model employed has high explanatory power for all of the regressions with R^2 values very close to 0.8. The regression outputs can be seen in detail in the Appendix, Tables 3 to 5.

6.3 First Sorting – UP and DOWN

The regression results for the UP and DOWN cases are presented in Table 3 in the Appendix. In the UP case, the intercept is positive but insignificant during the two year event period. By contrast, for the DOWN case the intercept is significantly negative. This suggests that fund managers of upgraded fund do not add value to investors net of transactions costs and fees, while downgraded funds generate negative abnormal returns to investors net of fees. These findings are consistent with those of Carhart (1997) and Fama and French (2008), whom, like us, find that after controlling for the Carhart (1997) four factors, fund managers generate insignificant or negative abnormal returns. Within the UP case, the intercept dummy is positive and significant at the ten percent level, suggesting that, other things being equal, fund managers improve performance in the year following a rating upgrade. For the DOWN case, the intercept dummy is negative, but not significant; suggesting that, other things being equal, a rating downgrade does not significantly affect performance. In light of the findings in the previous section that rating upgrades (downgrades) are associated with a subsequent inflow (outflow) of new capital, these findings suggest that the funds in our sample are able to benefit from economies of scale, contradicting the prediction of Berk and Green (2004).

We find that both up and downgraded funds have significant positive loadings on the excess market return and on SMB, both of which are slightly higher for upgraded funds during the event period than for downgraded funds. However, unlike downgraded funds, upgraded funds have significant positive loadings on MOM. Following a rating change, we find that upgraded funds reduce loadings on all factors, as is indicated by the negative factor dummies. However, none of these are significant at any higher levels. Still, the fact that all factor dummies are negative (and the p-values of the coefficient are not terribly high) lends some support to the idea that good fund managers reduce the riskiness of their portfolio in response to good performance to lock-in their ranking relative their peers. For downgraded funds, the picture is more mixed. The factor dummies show that loadings on the market excess return and SMB increase while loadings on HML and MOM decrease, although none of these changes are significant. These findings suggest that a rating downgrade does not have a significant effect on the strategy pursued by fund managers as a group.

Key Findings

- **Rating upgrades, other things being equal, lead to significant improvements in performance.**
- **Rating downgrades, other things being equal, do not significantly affect performance.**
- **Upgraded funds tend to decrease loadings on all factors.**

6.4 Second Sorting – Rating Change Type

The regression results of the rating change types are presented in Table 4 in the Appendix. The intercept for all cases experiencing a rating upgrade are positive except for case 12, but the only significant intercept is in case 45. The intercept for all cases experiencing a downgrade are negative, but only 32 and 21 are significant, which is not surprising since they had lower ratings to start with. We find positive intercept dummies for all upgraded cases except case 45, but only the intercept dummy for case 34 is significant. This suggests that these fund managers, other things being equal, improve performance in the year following a rating change. These funds also experienced a significant increase in their TNA following their rating upgrade, lending support to the idea that these funds benefit from economies of scale. Although the intercept dummy is not highly significant for case 12, it is not highly insignificant either (p-value of 13 percent). This is interesting in light of the findings in Section 6.1 that these funds experienced a significant decrease in size, which according to the Berk and Green (2004) framework should result in improved performance. For all cases experiencing a rating downgrade, the intercept dummy is negative but insignificant; suggesting that fund performance is not significantly affected by a rating downgrade.

For all cases in the rating change type sorting, funds have loadings of approximately 1 on the excess market return. The only exceptions being cases 54 and 45, which have considerably lower loadings on the market excess return compared to the other cases. This is interesting in light of the risk premia for the market excess return being negative for the sample period, which could account, in part, for these funds having the highest ratings. Almost all rating change type cases have significant loadings on SMB, where once again 5 star funds that are downgraded stand out, as it is the case with the highest loading on SMB. Upgraded funds also tend to have higher positive loadings on the MOM factor compared to downgraded funds, which could be a contributing factor to their upgrades (downgrades). By contrast, downgraded funds to a larger extent load on the HML factor. The results show that of the thirty-two factor dummies only five of them are significant, four of which belong to downgraded funds, suggesting that downgraded funds are more responsive to rating changes than upgraded funds, but not to any greater extent. However, the factor dummies, both significant and insignificant, for the downgraded funds have different signs, which in aggregate cancel out. Consequently, as a group, no systematic change in risk is observed for downgraded funds. Among the cases, the change in loadings is most pronounced in 5 star funds that are downgraded, which significantly increase loadings on the excess market factor and decrease loadings on SMB. One explanation for this could be that fund managers that have been top-ranked and subsequently lose their position feel a strong sense of loss of prestige, which induces them to alter their strategy to a larger extent compared to other downgraded funds. 3 star funds that are downgraded significantly

decrease loadings on HML while 2 star funds that are downgraded significantly increase their loading on HML. Overall, these findings seem to suggest that rating downgrades, other things being equal, influence the investment strategy of fund managers more than rating upgrades, for certain rating change types, although not to any greater extent. These effects are rather small and random for the sample as a whole and are largely limited to 5 star funds that are downgraded.

Key Findings

- **Downgraded funds tend to be more responsive to rating changes, than upgraded funds. But because the direction of the changes in loadings are different, they cancel out in aggregate.**
- **5 star funds that are downgraded seem more inclined to alter their strategy in the year following their rating change than upgraded funds.**

6.5 Third Sorting – TNA

The regression results for the fund size quintiles are presented in Table 5 in the Appendix. For upgraded quintiles, the intercept is close to zero except for the largest quintile, which has a significantly positive intercept, suggesting that large funds on average are better at generating higher abnormal returns. All downgraded quintiles have negative intercepts, which are significant for the two smallest quintiles. Again, this suggests that larger funds perform better than smaller funds. These findings are contrary to Fama and French (2008) who find that small funds on average have higher alphas than large funds. However, it should be noted that we do not control for other factors that may impact on performance. For example, Carhart (1997) finds that expense ratios and turnover significantly reduce the abnormal returns realized by investors, something we do not control for. Following a rating upgrade, all quintiles except the largest experience an improvement in performance as is indicated by the positive intercept dummies. Interestingly, the largest quintile is the only one of the upgraded cases with an average TNA greater than $3.5 billion which Latzko (1999) finds is somewhat of an upper limit for achieving economies of scale. However, it should be noted that the change is highly insignificant. Although upgraded funds as a group improve performance significantly following a rating change, this effect is not present in all fund size cases as only UP4Q experienced a significant improvement in performance. For downgraded funds, the intercept dummies are negative but it is only significant for DOWN3Q.

There is little dispersion across the different quintiles as loadings on the market excess return, for both up and downgraded, are very close to 1 and are all highly significant. A few things are worth taking note of. For the upgraded funds, all quintiles have significant loadings on SMB except the largest one. This is not surprising in light of the theoretical discussion that large funds are likely to face greater liquidity

constraints preventing them from investing in the smallest and most illiquid stocks. Further, for upgraded funds, the smallest quintile had significant loadings on all four factors whereas the largest only loaded significantly on the market excess return and HML. This supports the idea proposed earlier that the largest funds to a larger extent follow a passive strategy attempting to track their benchmark whereas smaller funds are more inclined to deviate from the passive benchmark and chase abnormal return. But they do so unsuccessfully, as the largest quintile has higher abnormal return than the smallest. However, this is not the case for downgraded quintiles. Among downgraded funds, the largest quintiles have significant loadings on as many factors as the intermediate quintiles. As in the case for upgraded funds, the largest downgraded funds do not load significantly on SMB while intermediate quintiles do. Somewhat surprising is that among downgraded funds, the smallest quintile does not load significantly on SMB, although this is probably more likely to be due to an active choice rather than liquidity constraints. Small and large funds being downgraded also seem equally likely to load on the MOM factor. Following a rating upgrade, fund managers tended to decrease loadings on all factors, however, only three out of twenty factor dummies were significant for the quintiles and they were not confined to a certain quintile. This suggests that the finding that upgraded funds tend to reduce risk is not affected by their size. Downgraded funds increased/decreased loadings to the same extent but only four out of twenty dummy factors were significant, three of which were negative and distributed evenly across quintiles. Overall, these findings suggest that funds of different size do not respond significantly differently to changes in their rating in terms of altering their strategies.

Key Finding

- **Funds of different size do not seem to respond differently, in terms of strategy, to rating changes.**

6.6 Results in Relation to Hypotheses Two to Six

In the year following a rating upgrade, other things being equal, fund managers on average see performance improve, as the significantly positive intercept dummy for the UP case indicates. However, upon closer inspection the observed improvement is confined to case 34. We do not find support that downgraded funds see performance improve as the intercept dummies are negative but insignificant, for sortings on both rating change type and fund size (except for case DOWN3Q). With regard to case 12, the positive intercept dummy is interesting because as was pointed out earlier, this case did not experience the expected increase in TNA but instead experienced a significant decrease in TNA. Within the Berk and Green (2004) framework, such a decrease should be associated with an improvement in performance, as we find, because of diseconomies of scale in asset management. Overall however, we find stronger

support for economies of scale in our study than for diseconomies of scale. These findings lend some support to reject Hypothesis 2 that abnormal returns decrease (increase) for funds receiving a higher (lower) rating.

For the different rating change types, performance only improved significantly for case 34. For the cases sorted on TNA, we found that the only significant changes in performance were confined to the intermediate quintiles, lending some support to the idea that the fund size – performance relationship may be non-linear. Although the Hypothesis 2 should be seen as rejected, we do find some support for Hypothesis 3 that funds with different initial ratings and sizes are affected differently, in terms of performance, by Morningstar rating changes even though the effect was the opposite of what we expected. This lends some support to Hypothesis 3 that changes in performance following a rating change vary across cases sorted on fund size and rating change type.

As a group, downgraded funds were unresponsive to their rating change in terms of altering loadings on factor which were random and insignificant, and individually they seem to alter loadings differently. We therefore find no support for Hypothesis 4 that loadings on factors increase for downgraded funds.

As a group, upgraded funds reduce their loadings on all factors, but insignificantly. This lends some support to Hypothesis 5 that loadings on factors change in response to a rating upgrade. These findings are consistent with those of Brown, Harlow and Starks (1996) that fund managers decrease the riskiness of their portfolio in response to good performance to lock-in their ranking relative their peers.

For the sortings on rating change type and size we found some significant factor dummies, but without any obvious pattern. This thesis therefore finds little support for hypothesis 6 that smaller and intermediate sized funds are more inclined to change their loadings on factors following a rating change compared to larger funds.

6.7 Limitations of the Results

In the sample selection procedure, only funds that experienced a rating change and were able to retain the new rating for at least one year have been included. Because rating upgrades and downgrades have different (potentially offsetting) effects, studying funds that have more than one rating change in its estimation window would not yield interesting results because the effects caused by upgrades/downgrades would be netted against each other so the regressions would only capture the aggregate effect, which would not be useful as the purpose of this study is to examine these effects in isolation. As a by-product of this methodology, however, our sample may be subjected to the look-ahead bias in the sense discussed in Section 5.2. Although we cannot be sure of the extent of this bias, our results suggest that these

concerns were not materialized, as there was some affect on performance but not the extent implied by a severe look-ahead bias.

Another issue relates to the representativeness of the data sample, which was discussed in the data section. As the average fund changes ratings 1.5 times per year, the sample used in this thesis might not be representative, which could result in a potential selection bias if the features of our data sample differ significantly from the population at large. However, in order to be able to study the effect of Morningstar rating changes, it is crucial that these effects be studied in isolation. Using shorter event windows would have been difficult as the effect of the rating changes on TNA would not have been fully realized. As with most regression studies there are compromises that need to be made to make the study feasible. However, we do not think that the extent of this selection bias is so great that inferences about the population are rendered impossible.

7. Conclusions

The purpose of this thesis was to examine whether Morningstar, when evaluating fund performance, in doing so, also affects those funds' performance and investment strategy. In this sense, Morningstar is not an impartial observer but an influencing institution. Theory provides compelling arguments that fund size and the incentives of fund managers to shift risk might be impacted by changes in Morningstar ratings which in turn drive changes in performance and strategy. We find support that even with the new rating system Morningstar still remains an influential source of information to investors who rebalance their fund portfolios, flocking to high performing funds and fleeing low performing funds, in response to their rating changes. This confirms the suspicion of Del Guercio and Tkac (2008) that their findings still apply under the new Morningstar rating methodology. Fund managers, like investors, also care about the Morningstar rating they hold and respond to changes in their rating by altering their portfolios. Particularly we find that fund managers that perform well, to some extent reduce the riskiness of their portfolio, lending support to the idea that they attempt to lock in their ranking relative their peers in response to their performance. Maybe fund managers care precisely because they know that investors care about the ratings they hold. This leaves the reader wondering if investors are right in basing their asset allocation decisions based on the Morningstar rating system? Given our findings that rating upgrades, other things being equal, are associated with an improvement in performance, the answer is yes. However, investors may not fully be able to benefit from this insight as at the time of the change investors cannot know whether funds will be able to retain their new rating. Further, the rating change – performance relationship is asymmetric as downgraded funds do not underperform to the same extent. Part of the explanation to the existence of this relationship may be because there are economies of scale in

asset management as upgraded funds, on average, see assets under management grow while downgraded funds see assets under management decrease. It should be stressed that the presence of economies of scale does not rule out the presence of diseconomies of scale. As we are unable to observe their effects in isolation only their joint effect, it could be that the economies dominate the diseconomies of scale in our sample. Disentangling the two may be a subject of future research.

8. Suggestions for Future Research

The interrelation of the impact of fund size on performance and strategy may be non-linear. It is of interest to study the impact of fund size on performance and strategy by first studying what the interrelation looks like. Specifically, it would be interesting to establish under what circumstances economies and diseconomies of scale are the most pronounced.

Perhaps our results lend little support to the Berk and Green (2004) framework because the time horizon of our event period is too narrow. Maybe the process predicted by their model that capital flows into and out of good (bad) funds continues for some time and equilibrium might only be reached in the longer term. If this is the case, the diseconomies of scale would also manifest themselves first over longer time horizons, which is interesting to look into.

It could also be of interest to replicate the study in this thesis but without isolating the rating changes. Meaning, by not using isolated events where the fund has to keep its rating for a certain amount of time. This entails using a model that is able to handle multiple rating changes within a short period of time. It might be difficult to get significant results, but the sample of funds would be highly representative.

If one had a much larger sample than the one used in this thesis, it would be of interest to study the Morningstar effect across rating change type cases and size at the same time. This would combine the theories presented in this thesis in an interesting way.

9. Reference List

9.1 Academic Sources

Adkisson, J. A., & Fraser, D. R. (2003). Reading the stars: Age bias in Morningstar ratings. *Financial Analysts Journal* , 24-30.

Ammann, M., & Verhofen, M. (2007). Prior Performance and Risk-Taking of Mutual Fund Managers: A Dynamic Bayesian Network Approach. *The Journal of Behavioral Finance* , 20.

Bechmann, K. L., & Rangvid, J. (2007). Rating mutual funds: Construction and information content of an investor-cost based rating of Danish mutual funds. *Journal of Empirical Finance* , 662-693.

Beckers, S., & Vaughan, G. (2001). Small is Beautiful. *Journal of Portfolio Management* , 9-18.

Berk, J. B., & Green, R. C. (2004). Mutual Fund Flows and Performance in Rational Markets. *The Journal of Political Economy* , 1269-1296 .

Blake, C. R., & Morey, M. R. (2000). Morningstar ratings and mutual fund performance. *Journal of Financial and Quantitative Analysis* , 451-484.

Brown, K., Harlow, V. W., & Starks, L. (1996). Of Tournaments and Temptation: An Analysis of Managerial Incentives in the Mutual Fund Industry. *Journal of Finance* , 85-110.

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

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance* , 57-83.

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* , 1276-1303.

Cochrane, J. H. (1999). New facts in finance. *Economic Perspectives, Federal Reserve Bank of Chicago* , 36-58.

Coleman, M. (2008, 02 27). *Seeking Alpha*. Retrieved 05 08, 2009, from seekingalpha.com: http://seekingalpha.com/article/66264-index-funds-continue-gaining-market-share-vs-active-management

Del Guercio, D., & Tkac, P. A. (2008). Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow. *Journal of Financial and Quantitative Analysis* , 907-957.

Fama, E. F., & French, K. R. (2008). Mutual Fund Performance. *Manuscript University of Chicago* , Section 1.

Fama, E. F., & French, K. R. (1993). The Common Risk Factors in the Returns on Stocks and Bond. *Journal of Financial Economics* , 3-56.

French, K. R. (2009). *Kenneth R. French*. Retrieved 04 30, 2009, from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Gerrans, P. (2004). Australian Managed Fund Ratings and Individual Investors. *Australian Journal of Management* , 87-108.

Goetzmann, W. N., & Peles, N. (1997). Cognitive dissonance and mutual fund investors. *The Journal of Financial Research* , 145-159.

Gottesman, A. A., & Morey, M. R. (2006). MORNINGSTAR MUTUAL FUND RATINGS REDUX. *The Journal of Investment Consulting* , 23-50.

Grinblatt, M., & Titman, S. (1989). Mutual fund performance : an analysis of quarterly portfolio holdings. *The Journal of Business* , 393-417.

Gujarati, D. N. (2003). *Basic Econometrics.* New York: McGraw-Hill/Irwin.

Hsiao, C. (2003). *Analysis of Panel Data.* Cambridge: Cambridge University Press.

Indro, D. C., Jiang, C. X., Hu, M. Y., & Lee, W. Y. (1999). Mutual fund performance: Does fund size matter? *Financial Analysts Journal* , 74-88.

Investment Company Institute. (2008). *Investment Company Institute.* Retrieved 03 30, 2009, from icifactbook.org: http://www.ici.org/stats/mf/2008_factbook.pdf

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance* , 65-91.

Jones, M., & Smythe, T. (2003). The Information Content of Mutual Fund Print Advertising. *The Journal of Consumer Affairs* , 22-41.

Kale, J. R., Hu, P., Pagani, M., & Subramanian, A. (2008). Fund Flows, Performance, Managerial Career Concerns, and Risk-Taking. *SSRN Working Paper Series* .

Keenan, C. (2002). Star turn. *Institutional Investor* , 124.

Knuutila, M., Puttonen, V., & Smythe, T. (2007). The effect of distribution channels on mutual fund flows. *Journal of Financial Services Marketing* , 88-96.

Latzko, D. A. (1999). Economies of Scale in Mutual Fund Administration. *The Journal of Financial Research* , 321-339.

McGuigan, P. (2006). THE STARS OF MORNINGSTAR: WHY THEY MAY NOT SHINE VERY BRIGHT. *Journal of Legal, Ethical and Regulatory Issues* , 95-108.

Morey, M. R. (2003). The Kiss of Death: A 5-Star Morningstar Mutual Fund Rating? *SSRN Working Paper Series* .

Morningstar. (2007, 07). *Morningstar*. Retrieved 04 30, 2009, from Morningstar.com: http://tw.morningstar.com/ODS/2000/Others/EU_CatDef_Jun07.pdf

Morningstar. (2008, 07 30). *Morningstar*. Retrieved 04 30, 2009, from Morningstar.com: http://corporate.morningstar.com/uk/documents/MethodologyDocuments/FactSheets/MorningstarRatingForFunds_FactSheet.pdf

Morningstar. (2009, 04 29). *Morningstar*. Retrieved 04 29, 2009, from Morningstar.com: http://corporate.morningstar.com/US/asp/subject.aspx?xmlfile=177.xml

Perold, A. F., & Salomon Jr., R. S. (1991). The Right Amount of Assets Under Management. *Financial Analysts Journal* , 31-39.

Pollet, J. M., & Wilson, M. (2008). How Does Size Affect Mutual Fund Behavior? *The Journal of Finance* , 2941-2990.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *The Journal of Finance* , 1589-1623.

Standard & Poors. (2009, 03 31). *Standard & Poors*. Retrieved 04 30, 2009, from Standardandpoors.com: http://www2.standardandpoors.com/spf/pdf/index/SP_500_Factsheet.pdf

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT press.

9.2 Non-academic Sources

Coleman, M. (2008, 02 27). *Seeking Alpha*. Retrieved 05 08, 2009, from seekingalpha.com: http://seekingalpha.com/article/66264-index-funds-continue-gaining-market-share-vs-active-management

Morningstar. (2007, 07). *Morningstar*. Retrieved 04 30, 2009, from Morningstar.com: http://tw.morningstar.com/ODS/2000/Others/EU_CatDef_Jun07.pdf

Morningstar. (2008, 07 30). *Morningstar*. Retrieved 04 30, 2009, from Morningstar.com: http://corporate.morningstar.com/uk/documents/MethodologyDocuments/FactSheets/MorningstarRatingForFunds_FactSheet.pdf

Morningstar. (2009, 04 29). *Morningstar*. Retrieved 04 29, 2009, from Morningstar.com: http://corporate.morningstar.com/US/asp/subject.aspx?xmlfile=177.xml

Standard & Poors. (2009, 03 31). *Standard & Poors*. Retrieved 04 30, 2009, from Standardandpoors.com: http://www2.standardandpoors.com/spf/pdf/index/SP_500_Factsheet.pdf

10. Appendix

Table 1 Change in TNA/S&P 500 following Morningstar rating changes

	UP	DOWN	12	23	34	45	54	43	32	21
t-statistic average	1.548	-2.298	-4.506	1.219	3.385	4.132	0.803	-0.927	-4.745	-4.011
p-value average	(0.069)	(0.017)	(0.000)	(0.119)	(0.001)	(0.000)	(0.216)	(0.183)	(0.000)	(0.000)
t-statistic median	1.830	-3.975	-6.597	1.223	3.873	6.275	-1.091	-1.943	-5.879	-5.045
p-value median	(0.042)	(0.000)	(0.000)	(0.118)	(0.001)	(0.000)	(0.144)	(0.034)	(0.000)	(0.000)
Observations	83	143	10	32	33	8	19	63	50	11

Table 2 Average TNA before and after rating change

USD Millions	UP	DOWN	12	23	34	45	54	43	32	21
Before	1 557	3 293	171	1 117	1 777	3 861	7 907	4 079	1 007	1 207
After	1 811	3 178	144	1 271	2 118	4 461	6 761	4 218	912	1 330

USD Millions	UP5Q	UP4Q	UP3Q	UP2Q	UP1Q	DOWN5Q	DOWN4Q	DOWN3Q	DOWN2Q	DOWN1Q
Before	5 785	1 280	473	177	64	14 600	1 566	535	243	79
After	6 808	1 448	547	195	80	13 834	1 536	508	214	73

Table 3 Regression results sorting UP and DOWN cases

	Intercept	Mkt-rf	SMB	HML	MOM	D intercept	D Mkt-rf	D SMB	D HML	D MOM	Sample size	R-square
UP	0.083	1.009 ***	0.213 ***	-0.013	0.064 ***	0.148 *	-0.033	-0.060	-0.054	-0.036	84	0.796
	(0.176)	(0.000)	(0.000)	(0.733)	(0.001)	(0.067)	(0.261)	(0.194)	(0.306)	(0.190)		
DOWN	-0.167 ***	0.964 ***	0.200 ***	-0.008	-0.016	-0.092	0.027	0.001	-0.049	-0.030	143	0.802
	(0.001)	(0.000)	(0.000)	(0.761)	(0.416)	(0.156)	(0.294)	(0.978)	(0.209)	(0.217)		

p-values within brackets *** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 4 Regression results sorting on rating change type cases

	Intercept	Mkt-rf	SMB	HML	MOM	D intercept	D Mkt-rf	D SMB	D HML	D MOM	Sample size	R-square
12	-0.275	0.923 ***	0.342 ***	-0.023	0.057	0.443	0.022	-0.081	0.163	-0.170	10	0.762
	(0.223)	(0.000)	(0.004)	(0.868)	(0.503)	(0.132)	(0.839)	(0.596)	(0.342)	(0.116)		
23	0.146	1.007 ***	0.297 ***	-0.099	0.072 *	0.022	-0.026	-0.052	-0.022	-0.045	32	0.798
	(0.126)	(0.000)	(0.000)	(0.147)	(0.075)	(0.866)	(0.628)	(0.498)	(0.808)	(0.350)		
34	0.047	1.042 ***	0.140 ***	0.045	0.057 **	0.237 **	-0.052	-0.056	-0.138 *	-0.017	33	0.835
	(0.604)	(0.000)	(0.003)	(0.389)	(0.014)	(0.040)	(0.193)	(0.408)	(0.067)	(0.652)		
45	0.483 ***	0.931 ***	0.106	0.143	0.086	-0.072	0.029	-0.130	-0.114	0.049	9	0.784
	(0.001)	(0.000)	(0.193)	(0.123)	(0.238)	(0.775)	(0.737)	(0.340)	(0.448)	(0.611)		
54	-0.000	0.812 ***	0.444 ***	0.079	-0.098	-0.197	0.217 ***	-0.269 **	-0.169	0.034	19	0.787
	(0.995)	(0.000)	(0.000)	(0.329)	(0.165)	(0.327)	(0.008)	(0.021)	(0.116)	(0.678)		
43	-0.119	0.985 ***	0.140 ***	-0.056	0.001	-0.064	-0.006	0.036	-0.005	-0.037	63	0.801
	(0.105)	(0.000)	(0.000)	(0.179)	(0.945)	(0.509)	(0.887)	(0.506)	(0.923)	(0.274)		
32	-0.245 ***	0.989 ***	0.189 ***	0.095 *	-0.014	-0.044	0.011	0.085	-0.140 **	-0.043	50	0.810
	(0.001)	(0.000)	(0.000)	(0.056)	(0.665)	(0.676)	(0.778)	(0.215)	(0.042)	(0.288)		
21	-0.387 ***	1.008 ***	0.171 **	-0.319 ***	0.034	-0.171	-0.036	-0.081	0.247 **	-0.026	11	0.834
	(0.010)	(0.000)	(0.024)	(0.001)	(0.556)	(0.401)	(0.652)	(0.479)	(0.045)	(0.730)		

p-values within brackets *** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 5 Regression results sorting based on size quintile cases

	Intercept	Mkt-rf	SMB	HML	MOM	D intercept	D Mkt-rf	D SMB	D HML	D MOM	Sample size	R-square
UP5Q	0.315 **	1.013 ***	0.058	0.115 *	0.057	-0.054	-0.032	-0.018	-0.182 *	0.038	16	0.862
	(0.021)	(0.000)	(0.313)	(0.056)	(0.152)	(0.724)	(0.558)	(0.836)	(0.056)	(0.494)		
UP4Q	-0.046	1.045 ***	0.209 ***	-0.026	0.041	0.451 **	-0.128 **	-0.084	-0.109	-0.005	17	0.747
	(0.777)	(0.000)	(0.006)	(0.795)	(0.427)	(0.027)	(0.048)	(0.436)	(0.440)	(0.938)		
UP3Q	0.032	1.027 ***	0.092 *	-0.005	0.031	0.027	-0.065	0.061	0.006	-0.020	17	0.837
	(0.792)	(0.000)	(0.097)	(0.938)	(0.521)	(0.862)	(0.302)	(0.451)	(0.954)	(0.727)		
UP2Q	-0.086	0.966 ***	0.390 ***	0.100	0.050	0.260	0.059	-0.161	-0.044	-0.132 **	17	0.815
	(0.547)	(0.000)	(0.000)	(0.263)	(0.117)	(0.177)	(0.370)	(0.135)	(0.687)	(0.033)		
UP1Q	0.116	0.996 ***	0.307 ***	-0.189 **	0.106 **	0.104	0.017	-0.096	0.016	-0.050	16	0.802
	(0.418)	(0.000)	(0.000)	(0.030)	(0.014)	(0.587)	(0.802)	(0.411)	(0.891)	(0.406)		
DOWN5Q	-0.055	0.946 ***	0.065	-0.117 ***	-0.076 ***	-0.120	0.055	-0.139 **	0.046	0.019	28	0.868
	(0.439)	(0.000)	(0.128)	(0.006)	(0.006)	(0.281)	(0.159)	(0.032)	(0.466)	(0.596)		
DOWN4Q	-0.118	0.996 ***	0.098 *	-0.031	0.026	-0.057	0.012	0.039	-0.094	-0.101 *	29	0.783
	(0.253)	(0.000)	(0.092)	(0.596)	(0.588)	(0.686)	(0.849)	(0.611)	(0.289)	(0.092)		
DOWN3Q	-0.085	0.906 ***	0.255 ***	0.048	-0.077 **	-0.280 **	0.164 ***	-0.087	-0.046	0.029	29	0.842
	(0.354)	(0.000)	(0.000)	(0.406)	(0.024)	(0.032)	(0.001)	(0.266)	(0.558)	(0.526)		
DOWN2Q	-0.267 *	0.981 ***	0.493 ***	0.051	-0.030	-0.016	-0.054	0.084	-0.074	0.006	29	0.764
	(0.093)	(0.000)	(0.000)	(0.571)	(0.677)	(0.931)	(0.490)	(0.438)	(0.484)	(0.940)		
DOWN1Q	-0.255 ***	0.986 ***	0.057	-0.079	0.103 ***	-0.125	-0.019	0.090	0.063	-0.138 **	28	0.830
	(0.004)	(0.000)	(0.224)	(0.180)	(0.003)	(0.280)	(0.697)	(0.190)	(0.399)	(0.003)		

p-values within brackets *** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 6 Breusch-Pagan and Haussmann test results

	UP	DOWN	12	23	34	45	54	43	32	21
Breusch-Pagan CHI-square statistic	7.80	0.30	1.29	1.34	11.37	3.09	0.01	0.56	0.22	1.73
Breusch-Pagan p-value	0.00	0.58	0.25	0.24	0.00	0.07	0.92	0.45	0.63	0.18
Haussmann test CHI-square statistic	13.34	26.6	1.35	4.04	4.17	14.95	26.00	11.09	10.00	N/A
Haussmann test p-value	0.27	0.00	0.99	0.96	0.96	0.18	0.00	0.43	0.53	N/A

	UP5Q	UP4Q	UP3Q	UP2Q	UP1Q	DOWN5Q	DOWN4Q	DOWN3Q	DOWN2Q	DOWN1Q
Breusch-Pagan CHI-square statistic	9.36	2.12	0.02	0.06	0.18	0.00	0.35	0.01	1.74	1.12
Breusch-Pagan p-value	0.00	0.14	0.89	0.80	0.67	0.94	0.55	0.90	0.18	0.28
Haussmann test CHI-square statistic	1.31	1.77	1.78	2.60	N/A	4.37	4.39	13.8	1.40	4.76
Haussmann test p-value	0.99	0.99	0.99	0.99	N/A	0.95	0.95	0.24	0.99	0.94

Table 7 Variance inflation factors (VIF)

	UP	DOWN	12	23	34	45	54	43	32	21
Mkt-rf	1.24	1.34	1.29	1.23	1.27	1.24	1.38	1.43	1.24	1.48
SMB	1.22	1.33	1.27	1.21	1.21	1.23	1.37	1.41	1.21	1.45
HML	1.04	1.07	1.08	1.10	1.09	1.10	1.08	1.08	1.08	1.07
MOM	1.09	1.02	1.05	1.08	1.02	1.05	1.02	1.01	1.05	1.01

	UP5Q	UP4Q	UP3Q	UP2Q	UP1Q	DOWN5Q	DOWN4Q	DOWN3Q	DOWN2Q	DOWN1Q
Mkt-rf	1.26	1.20	1.21	1.34	1.27	1.21	1.41	1.42	1.30	1.46
SMB	1.23	1.17	1.18	1.32	1.22	1.20	1.39	1.40	1.30	1.43
HML	1.09	1.10	1.09	1.09	1.13	1.09	1.05	1.09	1.07	1.09
MOM	1.04	1.06	1.07	1.01	1.06	1.06	1.01	1.01	1.02	1.02

Table 8 Wooldridge test for autocorrelation in panel data, H0: No first-order autocorrelation

	UP	DOWN	12	23	34	45	54	43	32	21
F-statistic	0.03	1.78	0.00	0.21	1.77	3.06	0.22	0.34	1.50	0.17
p-value	0.85	0.18	0.98	0.64	0.18	0.11	0.63	0.56	0.22	0.68

	UP5Q	UP4Q	UP3Q	UP2Q	UP1Q	DOWN5Q	DOWN4Q	DOWN3Q	DOWN2Q	DOWN1Q
F-statistic	1.57	0.39	2.18	2.17	1.89	0.57	0.24	0.54	1.58	0.09
p-value	0.22	0.53	0.15	0.15	0.18	0.45	0.62	0.46	0.21	0.76

Table 9 Likelihood ratio test for heteroskedasticity, H0: Homoskedastic

	UP	DOWN	12	23	34	45	54	43	32	21
Likelihood-ratio test CHI-square statistic	1028.04	1768.29	95.86	285.78	606.70	34.56	201.76	1047.96	548.76	74.47
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	UP5Q	UP4Q	UP3Q	UP2Q	UP1Q	DOWN5Q	DOWN4Q	DOWN3Q	DOWN2Q	DOWN1Q
Likelihood-ratio test CHI-square statistic	126.72	387.83	227.57	158.82	74.30	162.40	573.06	458.77	322.93	249.15
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 10 List of funds in the sample

300 North Capital Sm Cp Gr I
AIM Global Health Care A
AIM Real Estate A
AIM Small Cap Equity A
Alger Capital Appreciation A
Alger SmallCap Growth Institutional I
Allegiant Large Cap Core Equity A
Allegiant Large Cap Growth A
Allegiant Small Cap Growth A
Allianz OCC Growth A
American Century Growth Adv
American Century Utilities Adv
Ariel Appreciation
Atlantic Whitehall Growth
Barclays Global Investors S&P 500 St
BlackRock Equity Dividend A
BlackRock Fundamental Growth A
BlackRock Value Opportunities A
Bridges Investment
Bridgeway Blue-Chip 35 Index
Brown Capital Mgmt Balanced Instl
Brown Capital Mgmt Small Co Instl
California Investment Equity Income
Calvert Capital Accumulation A
Calvert New Vision Small Cap A
Chesapeake Growth A
Clipper
Columbia Marsico Growth A
Columbia Small Cap Core A
Columbia Small Cap Value II A
Delaware Aggressive Allocation A
DFA U.S. Large Company I
Dodge & Cox Stock
Dreyfus Basic S&P 500 Stock Index
Dreyfus Equity Growth A
Dreyfus MidCap Index
Dreyfus Midcap Value A
Dreyfus Small Cap Value A
Dreyfus/The Boston Company Sm Cp T/S Eq
Dreyfus/The Boston Company Small Cap Gr
Dryden Small Cap Core Equity A
DWS Blue Chip A
Eagle Mid Cap Growth A
Eagle Mid Cap Stock A
Eaton Vance Tax-Mgd Growth 1.0
Elfun Trusts
FBP Value
Federated Stock
Fidelity
Fidelity Advisor Financial Services A
Fidelity Advisor Growth & Income A
Fidelity Dividend Growth
Fidelity Equity-Income
Fidelity Equity-Income II
Fidelity Growth & Income
Fidelity Low-Priced Stock
Fidelity Select Banking
Fidelity Select Natural Resources
Fidelity Small Cap Independence
Fidelity Spartan Total Market Index Inv
Fidelity Value
Fifth Third Equity Index A
First American Equity Income A
First American Equity Index A
First American Small Cap Growth Opp A
First Investors Blue Chip A
Flex Funds Total Return Utilities
Franklin Balance Sheet Investment A
Gabelli ABC
GE Instl U.S. Equity Inv
GE Premier Growth Equity A
GE S&S Program Mutual
GMO U.S. Equity Allocation III
GMO U.S. Growth III
GMO U.S. Intrinsic Value III
Goldman Sachs Capital Growth A
Goldman Sachs Mid Cap Value A
Hartford Small Cap Value HLS IA
Hartford Small Company HLS IA
Heritage Diversified Growth B
Heritage Mid Cap Stock B
HighMark Value Momentum Fid
HSBC Investor Opportunity
ICON Healthcare
ICON Industrials
ICON Information Technology
ING DIRECT Index Plus Small Cap O
ING Index Plus SmallCap A
ING Small Cap Opportunities A
Integrity Growth & Income A
ISI Strategy
Ivy Small Cap Value A
Ivy Value A
Janus Aspen Forty Instl
Janus Global Life Sciences
Janus Twenty
Jennison Growth A
Jennison Mid Cap Growth A
Jennison Utility A
Jennison Value A
JHancock Sovereign Investors A
JHT Blue Chip Growth Trust Ser I
JPMorgan Equity Index A
Lazard U.S. Mid Cap Equity Instl
Legg Mason Partners Appreciation A
Legg Mason Partners Investors Value A
Longleaf Partners Small-Cap
Lord Abbett Growth Opportunities A
Lord Abbett Large-Cap Core A
Madison Mosaic Investors
MainStay Equity Index A
MainStay Mid Cap Value A
MainStay Small Cap Growth A
Mairs & Power Growth Inv

Managers Small Cap
Marshall Large-Cap Value A
Marshall Mid-Cap Value A
Marsico Focus
MassMutual Select Fundamental Value A
MEMBERS Large Cap Growth A
MEMBERS Large Cap Value A
MFS Utilities A
Morgan Stanley Equally-Wtd S&P 500 A
Morgan Stanley Inst Capital Growth I
Morgan Stanley Inst US Md Cp Value I
Morgan Stanley Special Growth A
Morgan Stanley Technology A
Morgan Stanley Technology B
Munder Internet A
Munder Large-Cap Growth A
Munder Mid-Cap Core Growth A
Mutual of America All America Instl
Nationwide A
Natixis AEW Real Estate A
Navellier Mid Cap Growth
Neuberger Berman Mid Cap Growth Adv
Nicholas II I
Nicholas-Applegate Growth Equity A
Nicholas-Applegate US Emerging Growth I
Nicholas-Applegate US Sys Lg Cp Gr I
Northeast Investors Growth
Northern Institutional Equity Index A
Northern Select Equity
Nuveen Multi-Manager Large-Cap Valu A
Oberweis Emerging Growth
Oppenheimer Capital Appreciation A
Oppenheimer Emerging Growth A
Oppenheimer Equity A
Osterweis
Perkins Mid Cap Value Instl
Pioneer Growth Leaders A
Pioneer Oak Ridge Small Cap Growth A
Pioneer Small Cap Value A
Principal MidCap S&P 400 Index Inst
Putnam Fund for Growth & Income A
Putnam Global Health Care A
Putnam Mid Cap Value A
Quaker Small-Cap Value A
Rice Hall James Micro Cap
RidgeWorth Mid-Cap Core Equity A
RiverSource Equity Value A
RiverSource Growth A
Robeco Boston Partners Sm Cap Val II Inv
RS Partners A
RS Select Growth A
RS Small Cap Core Equity A
RSI Retirement Trust Emerging Growth Eq
Schwab Small Cap Index Inv
Security Equity A
Security Large Cap Value A
Security Mid Cap Growth A
SEI Asset Allc Diversified Global Stk A
SEI Asset Allc Diversified U.S. Stock A
SEI Instl Mgd Large Cap Value A
Seligman Growth A
Seligman Large Cap Value A
Sentinel Small Company A
Sentinel Sustainable Core Opp A
Sequoia
Skyline Special Equities
SSgA S&P 500 Index
SSgA Tuckerman Active REIT
STAAR Larger Company Stock
Stratton Monthly Dividend REIT
SunAmerica Growth & Income A
T. Rowe Price Equity Index 500
T. Rowe Price Financial Services
T. Rowe Price Instl Large Cap Growth
T. Rowe Price Media & Telecommunications
T. Rowe Price Mid-Cap Growth
T. Rowe Price Small-Cap Stock
T. Rowe Price Tax-Efficient Growth
T. Rowe Price Total Equity Market Idx
Target Small Capitalization Growth
Thrivent Large Cap Value A
Transamerica Legg Mason Ptnrs Inv V A LW
Transamerica Partners Instl Small Core
UBS U.S. Large Cap Equity Y
UBS U.S. Large Cap Growth A
UBS U.S. Small Cap Growth A
Van Kampen Utility A
Vanguard Extended Market Idx
Vanguard Growth Equity
Vanguard Growth Index
Vanguard Institutional Index
Vanguard Selected Value
Vanguard Small Cap Index
Vanguard Strategic Equity
Vanguard Tax-Managed Growth & Inc
Vanguard Value Index
Vanguard Windsor II
Vantagepoint Broad Market Index I
Vantagepoint Growth & Income
Victory Special Value A
Virtus Growth Opportunities A
Virtus Index A
Wall Street
Wasatch Micro Cap
Wells Fargo Advantage Growth Equity A
Wells Fargo Advantage Opportunity A
Westport Select Cap I
White Oak Select Growth
Wilshire Large Company Value Instl
Wilshire Small Company Growth Instl