On the Behavior of Mutual Fund Investors and Managers

On the Behavior of Mutual Fund Investors and Managers

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit van Tilburg, op gezag van de rector magnificus, prof. dr. F.A. van der Duyn Schouten, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op vrijdag 22 november 2002 om 14.15 uur door

Alexei Petrovich Goriaev

geboren op 29 april 1977 te Neftekamsk, Rusland.

PROMOTORES:	Prof. dr. Th.E. Nijman
	Prof. dr. B.J.M. Werker
COPROMOTOR:	Dr. F.A. Palomino

Preface

Several chapters of this thesis were written in cooperation with others. Chapters 3 and 5 are joint work with Theo Nijman and Bas Werker, while Chapter 6 is co-authored with Frederic Palomino and Andrea Prat.

Acknowledgements

This thesis is the result of four years of work at the Department of Finance and CentER of Tilburg University, whose financial and organizational support is gratefully acknowledged.

First of all, I would like to thank my supervisors without whom this work would have never been accomplished. Theo Nijman introduced me to the topic of mutual funds and gave me indispensable insights in finance as well as general advice. Bas Werker was always available for inspiring discussions, in which he time after time corrected my imprecise statistical statements. Fred Palomino shared a lot of new ideas with me, encouraging me to look at things from a different perspective. I greatly appreciate their kindness and patience answering all my (silly) questions and reviewing numerous versions of the chapters in this thesis. Special thanks go to other members of the committee, Susan Christoffersen, Jenke ter Horst, Frank de Jong, and Chris Veld, for the time and effort they spent on my work.

Furthermore, I am indebted to all my colleagues and friends in Tilburg - impossible to mention them all - for creating a relaxed and stimulating atmosphere both during and outside working hours. Finally, I am grateful to my family - parents and older brothers, who gave me support in whatever direction I chose to pursue. May the spirit be with us.

Alexei Goriaev

Tilburg, June 2002

Contents

Chapter 1

Introduction

1.1 Motivation

Mutual funds have become one of the largest financial intermediaries in the leading world economies, currently controlling about 7 trillion dollars in assets in the US and over 3 trillion Euros in assets in Europe (see Investment Company Institute, 2002). Currently, investors can choose from thousands of funds offering a wide range of investment profiles, from relatively safe short-term debt instruments to relatively risky stocks and derivatives.

Similarly to investing in the stock market directly, holding mutual fund shares involves financial risks, as the fund's portfolio may rise or fall in value. Mutual funds claim to provide a number of benefits to their shareholders, compared to investing in other financial intermediaries or directly in the financial markets (see, e.g., Pozen, 1998):

1. *Low transaction costs.* Mutual funds allow investors including those with limited wealth to hold a diversified portfolio of financial securities at low cost. Mutual fund shares are easy to buy through an intermediary or directly, via telephone or Internet.

2. *Customer services.* Shareholders can transfer money between funds within the same family at low cost. In addition, they do not run liquidity risk, since they can sell their shares at net asset value at any time.

3. *Professional management.* The investment strategy of a mutual fund is developed by financial professionals, who are able to select the right stocks at the right time.

Thus, mutual funds claim to be especially attractive for small investors who do not have sufficient resources to follow a sound investment strategy at low cost.

Given the tremendous size of the mutual fund industry, it is crucial for the regulatory agencies to ensure that the funds efficiently invest money of their shareholders, since even a basis point difference in fund returns implies almost a billion dollar gain or loss for investors. The role of the academic research is to check the validity of the claims referred to above. It has been demonstrated that investing in mutual funds may not necessarily be optimal for consumers. It has been shown that active funds, on average, do not earn positive performance adjusted for risk and expenses (see, e.g., Gruber, 1996). Even though some funds seem to have superior risk-adjusted performance, there are many funds that consistently underperform their benchmarks (see, e.g., Carhart, 1997, and Kosowski et al., 2000). However, most shareholders of funds with consistently poor performance do not punish them by withdrawing their money, which may be due to various institutional and psychological factors (see, e.g., Gruber, 1996, and Sirri and Tufano, 1998). On the other hand, the concentration of money flows among a few top performers may provide adverse incentives to fund managers to take excessive risk in order to maximize the probability of becoming the top (see, e.g., Hvide, 1999, and Carpenter, 2000).

The aim of this thesis is to investigate empirically and theoretically the behavior of mutual fund investors and managers. These two problems are linked to each other, since in practice a manager's compensation is typically based on a proportion of the fund's assets (see, e.g., Khorana, 1996). In this thesis, we concentrate on those aspects of the allocation rules used by investors, which may provide adverse incentives to fund managers. On the other hand, we investigate strategies used by fund managers in response to these incentives.

1.2 The organization and structure of the mutual fund industry

In this section, we describe the organization and structure of the mutual fund industry, which is crucial for understanding the incentives and actual behavior of fund investors and managers. According to the basic definition, a mutual fund is an investment company that pools money from shareholders and invests in a diversified portfolio of securities (see, e.g., Investment Company Institute, 2002). In the US, the most important laws regulating mutual funds and ensuring investor protection are the Investment Company Act (ICA) and the Investment Advisers Act (IAA) of 1940. Mutual funds are typically organized as corporations and have a board of directors or trustees, which is elected by the shareholders. In contrast to most business corporations, mutual funds have very limited internal resources and rely on the provision of the specific services by affiliated organizations and independent contractors. In particular, the board of directors hires a separate entity - the investment advisor / management company - to provide all management and advisory services to a fund for a fee, which is usually based on a percentage of the fund's average net assets. In practice, however, the usual procedure is for the management organization to create mutual funds. To mitigate a potential conflict of interest, the ICA requires that an investment advisor must serve under a written contract approved initially by a vote of the shareholders and thereafter approved annually by the board of directors. Transactions between a fund and its manager are prohibited and at least 40% of a fund's directors must be independent from the fund's management company or principal underwriter. The IAA imposes recordkeeping, reporting, disclosure, and other requirements on investment advisors and contains several antifraud provisions. An investment advisor has a general fiduciary duty with respect to the compensation for its services, which bars an advisor from inadequate increase of its fees. Besides management company, mutual funds also employ principal underwriters who are responsible for the distribution of fund shares, custodians holding securities from fund portfolio, transfer agents conducting recordkeeping, and administrators overseeing the other agents providing services for a fund.

Mutual funds are considered "open-end" companies, since they are obliged to sell or redeem their shares at the net asset value (NAV), which is equal to fund's total net assets (total assets minus total liabilities) divided by the outstanding number of shares. The NAV must reflect the current market value of the securities in the fund portfolio and is usually calculated daily on the basis of the closing prices.

Mutual funds can be *active* pursuing their own portfolio management strategy or *passive* tracking the return of some benchmark index. In addition, mutual funds differ with respect to the share distribution method used. *Load* funds distribute their shares through broker-dealers who charge investors a commission proportional to the amount of the investment. Load fees may be front-end (charged at the time of the purchase) or back-end (charged at the time of the redemption). For the US funds, the front-end load is on average between 4% and 5%, while the back-end load usually declines the longer a shareholder holds the fund shares, e.g., from 5% after one year to 4% after 2 years, etc. (see, e.g., Pozen, 1998). In addition, brokers often receive annual distribution fees, called 12b-1 fees, typically ranging from 25 to 75 basis points of assets per year. *No-load* funds use direct distribution channels such as mail and phone and charge no front- or back-end loads and limited (up to 25 basis points per year) 12b-1 fees. Many funds have multiple share classes of the same fund corresponding to different combinations of load and 12b-1 fees. For example, class A shares are usually sold with a front-end load, while class B shares - with a back-end load. Besides the 12b-1 fees, the annual fund operating expenses paid by the shareholders also include the management fee, the recordkeeping fee, etc.

There are four basic types of mutual funds: equity, bond, hybrid, and money market (see Investment Company Institute, 2002). Equity and bond funds concentrate their investments in stocks and bonds, respectively. Hybrid funds typically invest in a combination of stocks, bonds, and other securities. These three types of funds are known as long-term funds, whereas money market funds are referred to as short-term funds, since they invest in securities maturing in less than one year. Morningstar, one of the leading mutual fund data providers, divides all long-term funds into four classes: domestic stock, international stock, taxable bond, and municipal bond.

1.3 Overview and contribution of the thesis

Chapter 2 of this thesis presents an overview of the main topics explored in the literature on mutual funds. Perhaps, the largest strand of this literature is devoted to the evaluation of mutual fund performance. Since the fund expected returns are affected by their risk exposures, the analysis is usually based on risk-adjusted performance measures. We discuss a number of studies that measure the average performance of mutual funds and examine factors explaining the differences in performance across funds. Another strand of the literature investigates the behavior of mutual fund investors, analyzing the impact of past performance and factors related to the transaction (in particular, information) costs on money flows to funds. Since managerial compensation is usually linked to the fund's size, the observed flow-performance relationship may provide adverse incentives to fund managers. In the third part of Chapter 2, we discuss the studies modelling the strategic response of fund managers to these incentives as well as empirical evidence on their actual behavior.

In Chapters 3 and 4, we present an empirical analysis of the behavior of mutual fund investors. Previous studies have identified a strong positive relationship between mutual fund flows and past performance (see, e.g., Gruber, 1996, and Sirri and Tufano, 1998). Most of these studies focus on the impact of the average past performance on fund flows at the annual frequency. In Chapter 3, we analyze the lag structure of the flow-performance relationship at the monthly frequency, using a sample of US growth funds in 1991-1999. We identify significant nonlinearities in the dynamics of investor reaction to past fund performance. In addition, we investigate whether investors pay more attention to raw rather than risk-adjusted performance of mutual funds.

Currently, there are several classification schemes that divide mutual funds into categories on the basis of the fund's stated investment objective or evaluated investment style. Different types of category-relative performance rankings are widely publicized in the media. Yet, little is known about the impact of performance relative to different classification schemes on mutual fund flows. So far, only the stated objective rankings have received attention in the literature and were found to be positively related to fund flows (see, e.g., Sirri and Tufano, 1998). In Chapter 4, we examine the relationship between flows to US mutual funds in 1993-1999 and their performance rankings within three types of categories: funds with the same stated objective, funds with the same Morningstar style, and funds within the same asset class. This allows us to learn which mutual fund classification schemes are used by investors to form peer groups for the evaluation of fund performance. In turn, this information is relevant for managers, who would like to know with which funds they should compete for investors' money. We also investigate whether cardinal or ordinal measures of fund performance (returns and return rankings, respectively) are more important for investors. In addition, we perform a category-specific analysis of the star spillover effect (see, e.g., Nanda, Wang, and Zheng, 2000), examining whether top performance of a star fund may be detrimental for flows to other funds in the same family.

Chapters 5 and 6 present the game-theoretic as well as empirical analysis of the behavior of mutual fund managers. In Chapter 5, we consider the statistical tests of risk taking by mutual fund managers performed in the literature. Several studies (see, e.g., Brown, Harlow, and Starks, 1996) report evidence in favor of the tournament hypothesis that within-year changes in risk are related to fund interim performance. However, Busse (2001) provides new evidence, based on daily data, which contradicts the previous evidence based on monthly data. Busse (2001) explains it by the fact that auto-correlation and cross-correlation in fund returns were not taken into account in the previous empirical tests of the tournament hypothesis. We contribute to this debate by considering the impact of both auto-correlation and cross-correlation on the tournament tests from an analytical point of view. First, we give analytical expressions for the biases arising in volatility estimates (based on both daily and monthly data) due to first-order autocorrelation effects in the daily fund returns. Second, to address the impact of cross-correlated fund returns on the tests, we provide explicit conditions under which the tests used in the literature have appropriate size properties.

In Chapter 6, we study risk taking incentives of mutual fund managers who have ranking objectives (as in a tournament). First, in a two-period model, we analyze the game played by two risk-neutral fund managers with ranking objectives. We show that in equilibrium, manager's choice of risk in the second period is negatively related to his relative performance over the first period. Using simulations, we also provide evidence that this result holds in the case with more than two competing funds. Second, we empirically test the predictions of the model in a sample of US diversified equity funds in 1980-1998. Specifically, we examine the relationship between fund choice of systematic risk in the last quarter of the year and relative performance over the first three quarters of the year.

Finally, Chapter 7 provides a summary of the main results in the thesis.

Chapter 2

A survey of the literature

2.1 Introduction

Mutual funds represent one of the organizational forms of delegated portfolio management, in which fund shareholders delegate the task of allocating their money to the fund manager. Since the manager's objectives are not necessarily identical to those of the fund's shareholders, a potential agency problem arises: the agent (fund manager) may not pursue investment policies optimal for the principals (fund shareholders). Numerous studies have examined the incentives and the actual behavior of mutual fund managers and investors. Among the main topics investigated in this literature are mutual fund performance evaluation, determinants of mutual fund flows, and strategic behavior of fund managers.

The measurement of mutual fund performance is crucial for evaluating fund managers. As discussed in Section 2.4.1, past performance of a mutual fund influences both the managerial compensation and the decision to retain, promote, or fire the manager. The central question in the studies of mutual fund performance is: "Does active fund management add value?" For a mean-variance investor, this question can be reformulated as: "Does the addition of active mutual funds to the portfolio of available assets lead to a shift in the mean-variance frontier?" If the answer is negative, consumers may be better off investing in low-cost index funds and avoiding expensively managed active funds. Two approaches have been used in the literature to measure risk-adjusted performance of mutual funds: return-based (see, e.g., Gruber, 1996) and portfolio-based (see, e.g., Daniel et al., 1997). The former approach employs fund returns, while the latter uses fund portfolio composition in order to construct a passive benchmark replicating the risk characteristics of the fund's portfolio. The difference between the fund's return and the benchmark return indicates whether the manager has superior knowledge or skills that allow him to outperform the benchmark (see Section 2.2.1). The existing empirical evidence suggests that mutual funds, on average, have a negative or, at best, neutral risk-adjusted performance (see Section 2.2.2). However, this does not necessarily imply that investors should not invest in mutual funds at all. Several studies examine whether there are consistent differences between performance of various mutual funds that can be forecasted (see Section 2.2.3). It has been found that there is a significant year-to-year persistence in raw returns, i.e., funds with the highest (lowest) raw returns over the last year are likely to be winners (losers) next year as well (see, e.g., Brown and Goetzmann, 1995). However, most of this persistence appears to be due to the differences in fund fees and exposures to the common risk factors (see, e.g., Carhart, 1997). Several studies nevertheless demonstrate that it is possible to identify funds with inferior as well as funds with superior risk-adjusted performance (see Kosowski et al., 2000) and that even investors with skeptical priors about the managerial skill may include the latter funds in their optimal portfolios (see, e.g., Baks, Metrick, and Wachter, 2001).

According to standard portfolio theory, an investor should base his allocation decision on the expected return and risk of mutual funds and alternative assets. Since in practice investors incur costs to collect and, maybe even more importantly, to process relevant information, they may limit their attention to a subset of the actual investment opportunity set, which does not necessarily include all mutual funds present in the market. Investors are more likely to consider more visible funds, for which the information or search costs are lower. Other factors related to the transaction costs, such as the fee structure (e.g., front load vs annual 12b1 fee), size of the fund family, and tax considerations, may also play a role for mutual fund investors. A number of studies investigates the relationship between performance and flows to mutual funds (see Section 2.3.2). Consistent with theoretical predictions, it has been demonstrated that better performing funds attract larger flows (see, e.g., Gruber, 1996). The flow-performance relationship appears to be convex, being stronger (weaker) for the best (worst) performers (see, e.g., Sirri and Tufano, 1998). The empirical evidence on other determinants of mutual fund flows is discussed in Section 2.3.3. Mutual fund flows are found to depend on a number of fund-specific factors, such as fund size, age, and fees (see, e.g., Sirri and

Tufano, 1998, and Chevalier and Ellison, 1997), as well as fund family characteristics, such as size and age of the fund's family and performance of other funds in the family (see, e.g., Nanda, Wang, and Zheng, 2000).

Numerous studies conduct a game-theoretic as well as empirical analysis of the strategic behavior of mutual fund managers (see Sections 2.4.2 and 2.4.3, respectively). There are two major factors that influence the expected payoff and, consequently, strategy of a mutual fund manager: the compensation structure and the retention policy. Several studies model the behavior of fund managers in response to the exogenously given compensation contracts observed in the mutual fund industry. They demonstrate that contracts linear or convex in the fund's benchmark-adjusted performance are not optimal for the incentive alignment between managers and investors (see, e.g., Admati and Pfleiderer, 1996). In equilibrium, fund managers typically choose lower effort and excessive risk taking (see, e.g., Hvide, 1999). In addition, the fund manager's risk policy may vary over time depending on the current performance relative to the benchmark (see, e.g., Carpenter, 2000). Some studies use a different approach allowing the compensation structure to be a part of the equilibrium, i.e., being endogenously determined in the model. They show that various types of fees used in the mutual fund industry may arise in equilibrium, including the incentive fee rewarding good performance (see, e.g., Das and Sundaram, 2002) and fraction-of-funds fee based on the fund's assets (see, e.g., Heinkel and Stoughton, 1994). The existing empirical evidence suggests that fund choice of risk may be related to its past performance (see, e.g., Brown, Harlow, and Starks, 1996, and Chevalier and Ellison, 1997). However, most of these results should be taken with caution, since they are based on statistical tests that do not take the auto-correlation and cross-correlation in fund returns into account (see Busse, 2001). Several studies find the evidence of the gaming behavior, such as window-dressing and marking-up of fund performance, by fund managers around the year-ends (see, e.g., Musto, 1999, and Carhart, et al., 2002).

2.2 Mutual fund performance evaluation

2.2.1 Definition of performance measures

In this section, we discuss the empirical evidence on mutual fund performance. We start by describing typical performance measures used in the literature. The most basic measure of mutual fund performance is a fund's raw return over a certain period of time. While being the simplest and most appealing to investors, this measure does not allow us to discriminate among managers who have superior skill, those who are lucky, and those who merely earn expected risk premiums on their high-risk investments. There are three factors driving mutual funds' expected raw returns: (i) the performance of the market and other risk factors, (ii) the fund's exposure to these risk factors, and (iii) the stockpicking skill of the portfolio manager. Various risk-adjusted performance measures have been constructed to single out the third factor, which plays an important role for investors choosing among funds and fund management companies devising managerial compensation. Most studies use absolute performance measures defined as a difference between the fund return and the return on a passive portfolio with a similar risk profile. The passive portfolio is formed using a return-based approach or a portfolio-based approach, which are explained below.

According to the return-based approach, fund performance is defined as the intercept in the time series regression of the excess fund return[1] on the excess returns of passive benchmark portfolios (factor-mimicking portfolios, in context of the arbitrage pricing theory):

$$R_{i,t} - R_t^f = \alpha_i + \sum_{k=1}^{K} \beta_i^k F_t^k + \varepsilon_{i,t}, \tag{2.1}$$

where $R_{i,t}$ is fund i's return, R_t^f is a risk-free rate, and F_t^k is the excess return on k-th benchmark portfolio in period t. This measure is often referred to as Jensen's alpha, since it was introduced in Jensen (1969), who used the excess market return as a single benchmark. Intuitively, Jensen's alpha can be interpreted as the difference between the fund's return and the return of the passive portfolio consisting of β_i^k units of the k-th benchmark ($k = 1, ..., K$) and $1 - \sum_{k=1}^{K} \beta_i^k$ units of the risk-free asset. A positive Jensen's alpha implies that mean-variance investors who used to restrict attention to the K benchmark assets and a riskless asset only, are able to extend their efficient set by

[1] Henceforth, the excess return denotes the rate of return in excess of the riskless interest rate.

taking a long position in the given fund, neglecting other effects such as the transaction costs and taxes.

Currently, most studies use multi-factor models to estimate Jensen's alpha. One of the most frequently used specifications is a three-factor model of Fama and French (1993). Besides an overall market factor, they use two additional stock market factors related to firm size (stock price times the number of shares) and book-to-market equity (the ratio of the book value of the firm's common stock to its market value). The corresponding factor returns are calculated as the difference between the returns on small- and big-stock portfolios and the returns on portfolios with high and low book-to-market equity, respectively. The four-factor model of Carhart (1997) adds one more factor related to one-year momentum in stock returns. The excess return on the corresponding factor-mimicking portfolio is computed as the difference between returns on stocks with high and low returns over the previous year. Thus, the Fama-French three-factor alpha measures fund performance taking into account exposure to size and growth factors, while the Carhart four-factor alpha in addition adjusts for the momentum effect.

In the portfolio-based approach, fund performance is measured as the difference between fund return and return on a passive portfolio with characteristics matching the portfolio of a fund under consideration. For example, Daniel et al. (1997) construct a synthetic portfolio of stocks matching fund holdings along the dimensions of size, book-to-market ratio, and one-year momentum. A zero performance measure indicates that the fund's performance could have been replicated by buying stocks with the same characteristics as those held by the fund, while a positive measure suggests that a manager has additional selection ability. In practice, funds are often assigned a stylized stock index as a benchmark, e.g., a small-cap index for funds investing in stocks of small companies. The simplicity of measuring fund performance as an index-adjusted return makes it appealing to investors. However, one should keep in mind that indexes based on relatively large market segments can provide only a rough approximation of the risk profile of a non-index fund. We will see that benchmarking by a certain index may change the investment strategy of the fund manager in a way detrimental for investors (see Section 2.4).

So far, we considered absolute performance measures calculated as the difference between the excess fund return and the return on the passive portfolio. Another type of absolute performance measure is the fund average excess return earned per unit of risk exposure. The most popular measure of this type is the Sharpe ratio, which is calculated as the average excess return of a fund divided by the standard deviation of the fund's returns:

$$Sharpe_i = \frac{\bar{R}_i - R^f}{\sigma_i}. \tag{2.2}$$

If the slope of the capital market line is larger than the fund's Sharpe ratio (the slope of the line connecting the position of the fund with the point of the risk-free rate), this is taken as evidence that the fund underperformed the market. Note that in contrast to Jensen's alpha, which takes the covariance of the fund return with benchmark returns into account, the Sharpe ratio is only based on the characteristics of a given fund. Therefore, the Sharpe ratio does not show whether an investor should add a given fund to his current portfolio, but helps to compare different mutual funds with each other. Specifically, a mean-variance investor restricted to invest either in fund A and a riskless asset or in fund B and a riskless asset will choose the one with the highest Sharpe ratio.

Absolute measures discussed above adjust fund performance for exposure to given passive benchmarks or risk factors. Another way to obtain a risk-adjusted performance measure is to evaluate fund performance relative to its peers, funds with a similar investment approach (i.e., funds with similar exposures to common risk factors). A typical relative cardinal measure of fund performance is the fund return in excess of the median or mean return in the fund's category. Note that this measure may not be appropriate if a fund's investment style differs significantly from those of other funds in the category. One should also keep in mind a potential effect of the survivorship bias, if the peer group contains only survived funds (as reported, e.g., by Brown and Goetzmann, 1995, disappearing funds tend to have poor performance). As shown in Section 2.4, the use of category-specific returns as a benchmark, similarly to benchmarking by stock indices, may lead to undesirable changes in fund strategies.

Most of the existing academic studies of mutual funds use cardinal performance measures as described above. However, the financial media as well as fund advertisements pay at least as much attention to ordinal performance measures based on the underlying cardinal measures. A typical ordinal measure is defined as a performance rank of a given fund within its category, which groups funds with a similar investment approach. The main difference between cardinal and ordinal performance measures is that the latter do not take into account by how much one fund outperforms the other. As discussed in Chapter 6, this can induce adverse risk-taking incentives to fund managers competing for the top performance ranks rather than maximizing risk-adjusted returns. Besides, ordinal performance measures are susceptible to the same criticisms as their underlying cardinal measures.

In Section 2.2.2, we describe the results of studies measuring the average performance of mutual funds, i.e., performance of the mutual fund universe taken as a whole. In Section 2.2.3, we discuss studies investigating whether there are consistent differences between performance of various mutual funds that can be forecast using various fund-specific and manager-specific characteristics.

2.2.2 Average performance of mutual funds

The existing empirical evidence based on both return-based and portfolio-based approaches suggests that an average active mutual fund has negative or neutral risk-adjusted performance net of expenses. This is demonstrated, for example, by Gruber (1996) whose main measure of performance is Jensen's alpha from a four-factor model with the market, size, growth, and bond factors. His sample consists of 270 US common stock funds during the period from 1985 to 1994 (almost all funds of this type that existed in 1984) and is free from survivorship bias. He finds that US stock funds underperformed an appropriately weighted average of the four benchmark indices by approximately 65 basis points per year. Since the average expense ratio in the sample is about 113 basis points per year, this implies that an average mutual fund earns positive risk-adjusted returns, but charges the investors more than the value added.

Similar conclusions are reached by Daniel et al. (1997) who measure performance of equity holdings of over 2500 US equity funds in 1975-1994 using a portfolio-based approach. They use as a benchmark the return on a portfolio of stocks that is matched to the fund's equity holdings each quarter on the basis of size, book-to-market, and one-year momentum characteristics. The authors find that US equity funds have some stock selection ability (i.e., buying those growth stocks that have higher expected returns than other growth stocks), but hardly any ability to time the different stock characteristics (i.e., buying growth stocks when they have unusually high returns). Overall, the performance earned by managers of active funds is not significantly greater than the difference between their expenses and expenses of passive index funds. Using the same sample of funds, Wermers (2000) extends this analysis by considering not only gross returns on funds' equity holdings, but also their net returns to investors. He finds that funds' stock portfolios outperformed the CRSP value-weighted market index by 1.3% per year, with 70 basis points being due to fund managers' stockpicking skills and the rest being due to the stocks' risk premiums. However, funds underpeformed the market index by 1% per year on a net return basis. The 2.3% difference between gross and net returns is due to the relatively low returns on fund nonstock holdings (0.7%), the expense ratios (0.8%), and the transaction costs (0.8%). Thus, a positive abnormal return earned by active mutual funds is more than offset by their expenses and transaction costs.

Ferson and Schadt (1996) criticize the standard approach to measure performance, which relies on unconditional expected returns. They argue that if expected returns and risks vary over time, then traditional performance measures may be upward- or downward-biased due to the common time variation in risks and risk premiums. They propose to use as a benchmark a managed portfolio strategy that can be replicated using publicly available information. Such conditional performance evaluation approach is consistent with the semi-strong form of market efficiency. In their model, Jensen's alpha is based on a factor model with time-varying conditional betas that are linear functions of the lagged public information variables including the short-term interest rate, dividend yield, term spread, and default spread. Using a sample of 67 US open funds from 1968 to 1990, they find that the distribution of the conditional Jensen's alphas is consistent with the neutral performance of mutual funds, whereas the unconditional Jensen's alphas indicate average underperformance.

Edelen (1999) argues that previously found negative performance of mutual funds may be explained by costs of providing liquidity to fund investors (open-end funds are obliged to buy and sell their shares at the net asset value). In his sample of 166 randomly selected open-end funds in 1985-1990, approximately one-half of the average fund's assets are redeemed in the course of the year and over two-thirds of the average fund's assets arrived as new inflow in the previous year. The author estimates that a unit of liquidity-motivated trading induced by investor flows, defined as an annual rate of trading equal to 100% of fund assets, is associated with 1.5-2% decline in risk-adjusted returns. Controlling for this liquidity cost changes the average Jensen's alpha from a statistically significant -1.6% per year to a statistically insignificant -0.2% per year.

2.2.3 Differential performance of mutual funds

In the previous section, it was demonstrated that mutual funds as a group have negative or neutral estimated performance adjusted for risk and expenses. However, this does not imply that consumers should avoid all mutual funds. If there exists a subset of funds that are able to consistently earn superior risk-adjusted returns, then investors would like to identify such funds and invest in them. In this section, we discuss the results of studies trying to identify consistent performance differences across funds and forecast fund performance.

Numerous studies examine whether past fund performance is indicative of future fund performance, i.e., whether there are differences in fund performance that persist over time. For instance, Brown and Goetzmann (1995) explore persistence in performance of US equity funds in 1976-1988 using both relative and absolute benchmarks. They find a significant year-to-year persistence in raw and risk-adjusted returns (the latter based on a three-factor model with the market, size, and bond factors) relative to the median return of all funds in the sample (relative benchmark) and S&P500 return (absolute benchmark). However, persistence seems to be mostly due to the underperforming funds. In other words, a fund underperforming other funds this year is likely to continue underperforming them next year. The authors note that the persistence pattern depends on the time period and that there was a significant reversal of relative winners and losers in a few years. They conclude that the observed pattern in relative performance could be due to the common component in fund strategies not captured by the standard risk-adjustment procedures.

This conclusion is supported by Carhart (1997) who demonstrates that most of performance persistence found in the previous studies can be attributed to the one-year momentum effect. His database covers US diversified equity funds in 1962-1993 and is free of survivor bias. When he sorts funds on the basis of lagged one-year raw return, his four-factor model with the market, size, book-to-market, and one-year momentum factors explains almost all of the cross-sectional variation in expected returns. In accordance with the previous evidence, funds with better last-year performance have higher return and one-factor Jensen's alpha than funds that underperformed last year. However, this difference is mostly due to the size and especially momentum factors, as last-year winners tend to hold more small stocks and momentum stocks than last-year losers. The only significant persistence unexplained by the Carhart's model is consistent underperformance by the worst-performing funds, which have significantly negative four-factor alphas. Investigating the factors explaining the differences in fund risk-adjusted performance, Carhart finds a significantly negative relationship between fund four-factor alphas and expense ratios, turnover, and load fees. A 1% increase in expense ratio, turnover, and maximum load fee is associated with 1.54%, 0.95%, and 0.11% decline in annual risk-adjusted return, respectively. Testing the consistency in funds' annual return rankings, Carhart finds that year-to-year rankings of most funds are largely random. Only funds in the top and bottom performance deciles in the last year are likely to remain in these deciles next year. As a result, one-year performance persistence is short-lived, being mostly eliminated after one year. Carhart finds slight evidence of persistence in risk-adjusted performance, as funds with high four-factor alphas tend to have above-average alphas in subsequent periods. However, this result should be taken with caution, since using the same model to sort and estimate performance may pick up the model bias that appears between ranking and formation periods.

Teo and Woo (2001) examine persistence in style-adjusted fund returns (fund returns in excess of the returns of the average fund in their Morningstar style category). They argue that most funds with high raw returns are clustered into well-performing styles and that a large year-to-year variation in style returns may preclude finding persistence in raw returns. Sorting funds on the basis of lagged three-year style-adjusted returns, they find significant spreads between Carhart's four-factor Jensen's alphas of funds from top and bottom deciles. These spreads are larger than those based on raw returns and persist for up to six years. This evidence suggests that some managers do have better abilities than the others.

Several studies investigate other factors that can explain mutual fund performance. Using a sample of US stock and bond funds in 1990-1999, Elton, Gruber, and Blake (2002) examine performance differences between funds using incentive fees (fees dependent on the fund's benchmark-adjusted return) and other funds using solely fraction-of-funds fees (fees proportional to the fund's assets). They find that funds with incentive fees earn, on average, an (insignificantly) positive multi-factor alpha of 58 basis points per year, which is higher than average alpha of other funds. Note, however, that this difference appears to be almost entirely due to differential expenses of these two classes of funds. Funds using incentive fees have an average expense ratio of 56 basis points per year lower than expense ratios of similar funds with no incentive fees. Among funds with incentive fees, the risk-adjusted performance seems to be higher when managers are hired internally by the fund family.

Chevalier and Ellison (1999a) study the relationship between fund performance and characteristics of fund managers that may indicate ability, knowledge, or effort. Their sample consists of 492 managers of growth and growth-and-income funds in 1988-1994. They find significant differences between raw returns of fund managers with different characteristics including the manager's age, the average SAT score at the manager's undergraduate institution, and whether the manager has an MBA. However, most of these return differences are attributed to the differences in managers' investment styles and to the selection biases. After adjusting for these, the authors find that managers who attended higher-SAT undergraduate institutions have higher risk-adjusted performance.

The beliefs of investors manifested in money flows to mutual funds also seem to contain some information about future fund performance. Gruber (1996) finds that US stock funds receiving more money subsequently perform significantly better than funds losing money. Using a sample of US equity funds in 1970-1993, Zheng (1999) shows that this "smart money" effect is short-lived and is largely but not completely explained by investors chasing past winners. She demonstrates that the smart money effect is not due to macroeconomic information or style effect, which suggests that investors use fund-specific information when choosing between funds. The smart money effect is mostly pronounced in the subset of small funds, whose lagged flows may be used to form the strategy beating the market.

Several studies use a Bayesian approach for performance evaluation, which combines prior investors' beliefs about the fund performance with the information in the data and produces posterior distribution of fund alphas. Baks, Metrick, and Wachter (2001) show that even some extremely skeptical priors about the skill of fund managers lead to economically significant allocations to some active diversified equity funds, based on posterior expectation of the Fama-French (1993) three-factor alpha. Pastor and Stambaugh (2002) develop a framework in which investors' prior beliefs can distinguish managerial skill from inaccuracy of the pricing model (CAPM, three-factor model of Fama-French, 1993, and four-factor model of Carhart, 1997). Using a sample of US domestic equity funds, they demonstrate that optimal portfolios of mutual funds are influenced substantially by both types of prior beliefs. Portfolios with the highest Sharpe ratios are constructed when prior beliefs have some confidence in a pricing model. However, investing in equity funds may be optimal even for skeptical investors who rule out the accuracy of pricing models as well as managerial skill.

Even if a small group of "star" fund managers earned superior risk-adjusted performance in the past, this may be due to luck. It is natural to expect that some funds out of thousands in the mutual fund universe outperform market indexes simply by chance. Using a sample of US equity funds in 1975-1994, Kosowski et al. (2000) apply a bootstrap technique to simulate the distribution of the extreme (maximum and minimum) performance measures across funds. Using various unconditional and conditional multi-factor models to measure performance, they demonstrate that the performance of the best and worst funds is not a result of sampling variability. To illustrate this point, 41 funds had a risk-adjusted return of at least 1% in 1995, while only 15 funds were expected to achieve this level by chance. This finding provides strong evidence of differential stockpicking skills among fund managers and supports the value of the active mutual fund management.

2.3 Behavior of mutual fund investors

2.3.1 Modelling mutual fund flows

In this section, we review studies conducting an empirical analysis of the determinants of mutual fund flows, focusing on the impact of past performance. In a typical regression model, the dependent variable is the fund's net relative or absolute flow. Traditionally (see, e.g., Gruber, 1996), net absolute flows are defined as the change in fund assets net of reinvested dividends:

$$F_{i,t} = TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t}), \tag{2.3}$$

where $TNA_{i,t}$ denotes fund i's total net assets at the end of period t and $R_{i,t}$ is return of fund i in period t. Similarly, net relative flows are defined as a net percentage growth of fund assets:

$$f_{i,t} = \frac{TNA_{i,t} - (1 + R_{i,t})TNA_{i,t-1}}{TNA_{i,t-1}} = \frac{F_{i,t}}{TNA_{i,t-1}}. \tag{2.4}$$

Both definitions are based on an assumption that all investor earnings are automatically reinvested in the fund and flows occur at the end of period t. A typical model in the literature specifies flows (in this case, net relative flows) as a linear function of past performance and a set of control variables:

$$f_{i,t} = a + b_1 r_{i,t-1} + ... + b_K r_{i,t-K} + x'_{i,t-1} c + u_{i,t}, \tag{2.5}$$

where $r_{i,t}$ is some measure of fund i's performance (e.g., raw return, Jensen's alpha, or corresponding ranking) in period t and $x_{i,t-1}$ includes such variables as fund size, age, fees, a measure of riskiness, and performance of other funds in the family. To control for unobserved individual effects (e.g., marketing effort, general reputation, etc.), $x_{i,t-1}$ sometimes includes lagged flow $f_{i,t-1}$.

The empirical evidence on the impact of past performance and other attributes of mutual funds on their flows is described in Sections 2.3.2 and 2.3.3, respectively.

2.3.2 Impact of past performance on mutual fund flows

The existing evidence demonstrates a strong positive relationship between mutual fund flows and various measures of their past performance measured over the one-year, three-year,and five-year horizons, including Jensen's alpha and raw return (see, e.g., Gruber, 1996) and category return rankings (see, e.g., Sirri and Tufano, 1998). When taken together, both raw and risk-adjusted performance measures have significantly positive impact on flows, although the impact of the latter appears to be stronger (see, e.g., Gruber, 1996). This suggests that some investors are style timers choosing funds with high loadings of factors that performed well recently. Note, however, that these effects may be partially offset by the negative impact of fund total risk on flows (see, e.g., Barber, Odean, and Zheng, 2001). The sensitivity of flows to performance seems to decline with time, i.e., fund last-year performance is more important for investors than fund performance two or three years ago (see Sirri and Tufano, 1998).

The flow-performance relationship appears to be asymmetric, as flows to top performers are more sensitive to their performance than flows to poorly performing funds. Using a piecewise linear model in a sample of US growth funds in 1971-1990, Sirri and Tufano (1998) show that flows to funds in the top performance quintile in their objective category are strongly related to their last-year return rankings, whereas for other funds the relationship between flows and performance is weak. For an average fund, moving five percentiles among the top performing funds in the category is associated with 8.4% increase in annual relative flow, while a similar move in rankings among funds with bad or intermediate performance results in 0 to 1.4% increase in flows. Chevalier and Ellison (1997) use a semiparametric model to estimate the shape of the relationship between fund flows and last-year market-adjusted returns (fund returns in excess of the market return) in a sample of growth and growth-and-income funds in 1982-1992. They demonstrate that this shape differs considerably in the subsets of young and old funds (funds with age of up to 5 years and over 5 years, respectively). For young funds, the shape of the flow-performance relationship is quite steep and close to linear. A 1% rise in the market-adjusted return of an average young fund is associated with about 4% increase in the fund's annual relative flow. In contrast, the expected flows to old funds are less sensitive to their last-year performance and the flow-performance sensitivity has a generally convex shape. Old funds outperforming the market are expected to attract about 2.8% extra annual flows due to 1% rise in the market-adjusted return.

Since performance persistence is more pronounced among poor performers than among good performers (see, e.g., Carhart, 1997), one may expect that consumers respond stronger to low than high performance. The divergence between these expectations and the observed convexity of the flow-performance relationship can be explained by a number of institutional and psychological factors, which prevent large outflows from funds with bad past performance. Market frictions such as the presence of search costs, back-end load charges, tax considerations, and restrictions of the investment retirement plans increase the transaction costs of withdrawing money from the poorly performing funds, while status-quo bias (see Zeckhauser, Patel, and Hendricks, 1991) and cognitive dissonance bias (see Goetzmann and Peles, 1997) make investors ignore information about bad fund performance.

Capon, Fitzsimons, and Prince (1996) use a different approach to examine the allocation rules used by mutual fund investors. They conducted a survey of 3000 consumers investing in US mutual funds who were asked to rate on a five-point scale the importance of given information sources and selection criteria and describe their investment approach and demographic characteristics. The results of the survey demonstrate that investors consider performance-related variables as the most important information source (published performance rankings) and selection criterion (performance track record). At the same time, fund characteristics other than return and risk, such as advertising (as an information source) and fund manager reputation, fund family scope, and management fees (as selection criteria), are also important for consumers. The authors also find that mutual fund clientele consists of several groups considerably differing from each other in terms of demographic characteristics and investment behavior. These groups range from the well-informed investors to the naive ones who are ignorant of their fund investment style and load structure. Further discussion of the non-performance factors driving mutual fund flows is carried on in the next section .

2.3.3 Impact of other factors on mutual fund flows

When the information about mutual fund performance is costly, consumers incur search costs to make an allocation decision. Many investors, especially small ones, may choose to save on these costs and make a choice based on the available (incomplete) information. In this case, more visible funds, i.e., the ones which are heavily advertised and have an established reputation, are expected to attract larger money flows, irrespective of their performance. In addition, flows to these funds may be more sensitive to their performance, since the impact of advertising and established reputation should be even stronger when combined with good performance. Fund flows may be also affected by factors related to other types of transaction costs, such as the fee structure (e.g., front load vs annual 12b1 fee), tax considerations, and the size of the fund family.

One proxy for fund visibility is its size. Apparently, large funds spend more on advertising and are more likely to receive media attention. Indeed, money flows to mutual funds are recognized to be roughly proportional to fund size (see, e.g., Gruber, 1996). This is the reason why most studies use the fund's relative flow as a dependent variable in the regressions. However, the magnitude of relative flows declines with fund size, i.e., large funds tend to attract significantly smaller relative flows than small funds (see, e.g., Sirri and Tufano, 1998). Therefore, size effect must be taken into account both in regressions of absolute and relative flows.

The level of media coverage, which helps to lower search costs, is found to be positively related to fund flows. Sirri and Tufano (1998) show that growth funds whose names are referred to in the major newspapers and periodicals attract larger flows during the same year, while Jain and Wu (2000) find that flows are significantly larger for those equity funds that are advertised in the financial magazines.

Fund age may also serve as a proxy for investor awareness about the fund. In contrast to young funds, old funds have an established reputation, which may be good or bad depending on their performance realized in the past. Therefore, recent performance should be more informative for young funds that do not have such reputation. Indeed, as discussed in Section 2.3.2, Chevalier and Ellison (1997) find that flows to young funds are more sensitive to their last-year performance than flows to old funds.

The effect of fund fees on flows can be twofold. On the one hand, higher fees may lead to lower flows, as investors would like to maximize net-of-fee earnings. In addition, load funds and funds with higher expense ratios have worse performance than funds charging lower fees (see Carhart, 1997). On the other hand, higher 12b1 fee, which is a part of the expense ratio, is associated with larger marketing expenditures and may increase fund flows. The existing evidence is consistent with the presence of both effects. Using a sample of US diversified equity funds in 1970-1999, Barber, Odean, and Zheng (2001) find that a negative relationship between fund flows and total fees (composed from load fees and expense ratios) is due to the strong negative impact of load fees. However, they find no significant relation between fund flows and expense ratios and even a positive relation in a subset of large funds. These results also suggest that investors pay more attention to salient fees, like loads and commissions, than expense ratio. The effect of advertising on fund investors may also explain higher flow-performance sensitivity of high-fee funds found by Sirri and Tufano (1998).

Bergstresser and Poterba (2002) study the impact of personal taxation on the investment decisions of consumers who hold mutual fund shares in conventional taxable accounts (not in tax-deferred retirement saving plans). Their sample includes US domestic equity funds in 1993-1999. They find that funds delivering more heavily taxed returns (i.e., returns including more dividends or realized capital gains) attract lower flows than funds with similar pretax returns and lower tax burdens. The flows also appear to be lower for funds with larger stocks of unrealized capital gains (new shareholders of such funds may be taxed on future distributions of these capital gains).

The magnitude of the transaction costs incurred by a mutual fund investor is also related to the characteristics of the fund's family. Since investors are more likely to be aware about the brand name of large and old fund families, funds from these families are more visible. In addition, families offering a large number of funds with a wide range of investment styles decrease the transaction costs for investors who often switch between different types of funds (e.g., stock funds and money market funds). Therefore, funds from large, old, and diverse families are expected to attract higher flows. Indeed, Ivkovic (2000) finds in a sample of US stock and bond funds in 1991-1999 that funds belonging to larger families attract higher flows. Using a sample of all US open-end funds in 1979-1998, Khorana and Servaes (2001) demonstrate that families achieve larger market share when they have more prior experience, offer funds in a wider range of objectives, and use more distribution channels. Nanda, Wang, and Zheng (2000) explore the performance spillover effects within the family, using a sample of US diversified equity funds in 1992-1998. They find that the presence of a star performer (fund with the return within top 5% in its category) in the family helps to boost flows to the other funds in the family.

2.4 Strategic behavior of mutual fund managers

2.4.1 The objectives of fund managers

Similarly to other industries, there is a potential divergence of interests between shareholders and managers of mutual funds. The manager's strategy consists of two major choices: *effort*, which allows him to extend the investment opportunities set, and *risk*, i.e., a point at this set. If the principal (fund shareholders) could contract directly on actions (effort and risk), it would be possible to achieve a first-best outcome with properly structured agent's (fund manager's) incentives. Since in practice the manager's effort is not contractible (i.e., not verifiable by a third party such as a court), the moral hazard problem cannot be eliminated.

In a typical mutual fund, two factors influence the manager's expected payoff: the compensation structure and the retention policy. Currently, two types of compensation schemes are used by mutual funds: base or fraction-of-funds fee and incentive fee (the latter always used in combination with the base fee). The base fee is linked to the fund's size and is charged as a percentage of the average net assets during the year (see, e.g., Khorana, 1996). Deli (2002) reports that in US marginal asset-based fee rates are greater for small funds, funds from small families, equity funds (compared to debt funds), and international funds (compared to domestic funds). These differences are interpreted as being due to the economies of scale and the difficulty of monitoring the performance.

The incentive fee depends on the fund's performance relative to a certain benchmark.

The 1970 amendment to the Investment Company Act of 1940 requires the incentive fees of US mutual funds be of a "fulcrum" type. This means that the fee must be symmetric around the benchmark, i.e., the reward for outperformance must be the same as penalty for underperformance. Probably, this restriction is the reason why only a few US mutual funds use incentive fees. According to Elton, Gruber, and Blake (2002), these are mostly large funds accounting for less than 2% of the total number of funds in the industry, but controlling more than 10% of the total assets under management. The incentive fees can be of the linear or bonus type, being the linear or discrete step functions of the benchmark-adjusted fund return, respectively. In most cases, funds use linear incentive fees with a limit (both upper and lower) on the size of the incentive fee, so that the sum of the base and incentive fees cannot be negative. As a result, the incentive fee is usually a piecewise linear function of benchmark-adjusted performance (flat below the lower limit and above the upper limit and increasing between them), which is convex up to the upper fee limit. As was discussed in Section 2.3.2, the sensitivity of flows to performance is higher for well-performing funds than for poor performers (see, e.g., Sirri and Tufano, 1998). This implies that the base fee is a convex function of the fund's past performance. Thus, fund performance influences the manager's expected payoff in a convex manner directly, through the incentive fee (over some range), and indirectly, through the base fee and the observed flow-performance relationship.

Another factor which influences manager's strategy is the impact of his actions on the probability of terminating the contract. Several studies demonstrate that fund performance plays a crucial role for the decision to dismiss, retain, or promote the fund's manager. Khorana (1996) estimates that managers in the lowest performance decile are four times more likely to be replaced than managers in the top performance decile. Chevalier and Ellison (1999b) find that the termination of the contract is more performance-sensitive for young managers, who do not have an established reputation, than for old managers. For young managers, the probability of termination is a convex function of past performance (over most of the range), decreasing steeply with performance in case of negative excess returns and being rather insensitive to the differences in performance at positive excess return levels. The authors also find that considerable deviations of the fund's sector weightings and the level of the unsystematic risk from the mean values in the objective category increases the probability of manager's termination in case of poor performance, while increasing, although to a smaller extent, the probability of his promotion in case of good performance. Thus, the convexity of the manager's expected payoff with respect to the fund's past performance may be weakened due to the strong impact of the manager's poor performance on termination decision.

There is vast literature providing extensive game-theoretic analysis of the managerial behavior in response to different payoff structures (see Section 2.4.2). A number of empirical studies test the predictions concerning the managers' risk-taking behavior based on these models as well as other hypotheses (see Section 2.4.3).

2.4.2 Managers' strategies: game-theoretic analysis

In this section, we discuss the studies modelling the strategic behavior of mutual fund managers. The models of the delegated portfolio management in the mutual fund industry, in which the agent (fund manager) receives money from the principal (fund shareholder) to invest in financial markets, have their own specifics. Since there are much more investors than funds, fund managers have most of the bargaining power. As a consequence, fund managers and not investors are typically proposing the compensation contracts. Therefore, most models of mutual funds examine pooling equilibria in which all managers have one type of contract and signal their quality with performance or separating equilibria in which managers signal the differences in their abilities by offering different types of contracts.

One strand of this literature adopts a behavioral approach and examines the equilibrium behavior of fund managers in response to exogenously given compensation structures observed in the mutual fund industry. Another strand of the literature models both the actions of fund managers and the investment strategies used by fund investors. In this case, the compensation scheme is determined endogenously within the model. In all these studies, the manager's compensation is some (linear or convex) function of the fund's performance with respect to some benchmark, which can be absolute (e.g., the return on a market index such as S&P500) or relative (e.g., the best return among other funds). In the former case, the benchmark is exogenous and cannot be influenced by players' actions. In the latter case, the benchmark is determined endogenously in the equilibrium.

We start with the first strand of the literature and exogenous benchmarks. Grinblatt and Titman (1989) use option pricing theory to analyze the impact of convex option-like compensation schemes on risk-taking behavior of fund managers. They show that such schemes induce excessive risk taking from both informed and uninformed fund managers. Moreover, managers with superior information may select the same portfolio as the uninformed managers, if the performance fee can be hedged in the manager's personal portfolio. Carpenter (2000) models the dynamic investment problem of a risk-averse manager who is compensated with a call option on the managed assets with an exercise price equal to a benchmark return and who cannot hedge this position. She demonstrates that option-like compensation does not always lead to greater risk taking. The manager dynamically adjusts volatility in response to changes in the benchmark-adjusted return and may actually decrease risk if the option is in the money or if the evaluation date is far away. Chen and Pennacchi (1999) analyze in a continuous setting the impact of the fund's prior performance on the portfolio choice of a fund manager with convex benchmark-adjusted compensation. They show that funds with poor performance have an incentive to increase the tracking error with respect to the benchmark, which is however not equivalent to an increase in volatility. Admati and Pfleiderer (1996) show that even compensation contracts that are linear in benchmark-adjusted performance are not optimal with respect to efficient risk sharing and incentive alignment between managers and investors. In their model, an optimal outcome is achieved when compensation is only based on the total unadjusted return of the manager's portfolio.

Similar conclusions are reached by studies in which managers are rewarded on the basis of relative performance, i.e., when the benchmark is endogenous. Hvide (1999) models the one-period game between fund managers with the tournament reward structure, where only the top performer receives the bonus (resulting, e.g., from the money flows). In his model, managers choose not only effort, which determines the expected return, but also the riskiness of the portfolio. In the extreme case, when there are no limits to possible risk taking, the tournament breaks down, as managers choose zero effort and infinite risk in equilibrium. When risk-taking is limited, the tournament rewards induce excessive risk taking and lack of effort from fund managers. The author shows that the scheme with higher reward for modest rather than excellent performance may lead to less risky strategies. Palomino (2002) analyzes a different reward structure, in which the manager's payoff depends linearly on the difference between his return and some function of the returns of other funds (e.g., the mean return in the fund's category). He shows that even in case of linear relative performance objectives, managers choose overly risky strategies to outperform their competitors. Furthermore, there is an underacquisition of information in equilibrium.

Thus, neither linear, nor convex compensation contracts can optimally (in the first-best sense) align interests of managers and investors of mutual funds. What happens if we allow the compensation structure to be determined in the equilibrium? Heinkel and Stoughton (1994) consider the multiperiod relationship between risk-neutral investor and a pool of risk-neutral fund managers with different, but ex ante unknown abilities. They show that in the first period the investor induces most managers to sign the standard ("boilerplate") contract with little performance-based component (only few managers with exceptional ability choose a different contract with high performance-based component). The investor provides proper effort-exerting incentives to fund managers by a credible threat of dismissal following a performance evaluation. The manager is only retained, if the return on his portfolio exceeds the benchmark by an appropriate amount (too high return indicates luck rather than skill). These results may provide theoretical justification for the limited use of performance fees in the mutual fund industry.

Huddart (1999) examines a similar two-period model with two risk-averse managers of different abilities. In this model, investors also infer about managers' abilities on the basis of their relative performance over the first period. In the second period, investors reallocate their wealth to the fund with the highest first-period return, which is most likely to be informed in equilibrium. However, this allocation rule, which maximizes investor perceptions of managerial ability, does not provide proper risk taking incentives to fund managers. When managers receive a fraction-of-funds fee, they choose overly risky strategies to maximize the chance of becoming the top after the first period. The uninformed manager does it to appear informed, while the informed manager does it to increase the cost of mimicking him. The author shows that the adoption of a performance fee with respect to an exogenous benchmark helps to mitigate these effects.

Das and Sundaram (2002) consider the setting in which fund managers choose fee structures to signal their abilities to investors and compare the equilibria with asymmetric incentive fees with the equilibria with (unlimited) fulcrum fees. Consistent with the previous studies, they show that asymmetric incentive fees encourage the adoption of more risky portfolios than fulcrum fees. However, when the entry costs for the uninformed managers are low, the incentive fees may be preferable for investors' welfare than fulcrum fees.

Palomino and Uhlig (2002) model a game in which risk-neutral investors choose between an index fund and an active fund. The manager of an active fund may be good or bad (a bad manager is uninformed, while a good manager may be informed with some probability) and is compensated with a fraction-of-funds fee. Investors can only observe realized returns, from which they infer about the unknown quality of the active fund's manager. Under the condition that investing in an active fund is not optimal ex ante (i.e., before observing returns), the model has an equilibrium, in which investing in the active fund is optimal ex post, if its return falls within some interval (i.e., is neither too low or too high). In this equilibrium, an informed manager picks a portfolio with minimal riskiness, and an uninformed manager chooses higher risk, gambling on a lucky outcome. When the fee structure is endogenous, both types of the active fund's manager choose the same fraction-of-funds fee structure.

2.4.3 Managers' strategies: empirical evidence

In this section, we review empirical evidence on strategic behavior of mutual fund managers. We start with the studies testing predictions of the theoretical models discussed in the previous section. Since the calendar year is often used as the performance evaluation period for mutual fund managers[2], they are interested in maximizing their calendar-year performance. The convexity of the manager's payoff in fund performance (see Section 2.4.1) suggests that mutual funds participate in the annual tournaments competing for the top year-end rankings. Based on the theoretical models of Carpenter (2000)and Chen and Pennacchi (1999), one can formulate the hypothesis that funds with bad performance after the first part of the year have an incentive to increase risk in the second part of the year, trying to catch up with interim winners at the end of the year. Several studies test this tournament hypothesis examining within-year changes in risk measured on the basis of monthly return data. Applying a contingency table methodology to the sample of US growth funds in 1976-1991, Brown, Harlow, and Starks (1996) find that

[2]In general, two types of the evaluation horizon are used: rolling horizon or fixed calendar-year horizon. Mutual fund performance based on the rolling one-year horizon (e.g., fund raw return during the last 12 months) as well as year-to-date performance (e.g., fund raw return from January to the current month) are often published in the financial newspapers. The calendar-year performance is reported in funds' prospectuses as well as fund listings published on an annual basis by many periodicals and data providers.

interim losers (defined as funds below the median category return over the first part of the year) increase risk towards the end of the year relative to interim winners. Using a sample of US domestic equity funds in 1992-1994, Koski and Pontiff (1999) apply regression methodology and find a negative relationship between fund return over the first semester and the change in total, systematic, and unsystematic risk between the first and second semesters. Chevalier and Ellison (1997) use a different approach, measuring fund risk on the basis of the fund's portfolio holdings. They also find a negative relationship between fund return over the first nine months of the year and the change in fund risk between September and December, using a sample of growth and growth-and-income funds in 1982-1992. However, Busse (2001) finds no such evidence, applying either the contingency table or the regression methodology to daily returns of 230 US domestic equity funds in 1985-1995 (new entrants after 1984 are not included). He explains this divergence in the results by the presence of the auto-correlation and cross-correlation in fund returns, which was not accounted for in the standard statistical tests used in the previous studies.

A related literature examines strategic changes in fund styles measured as factor loadings from a multi-factor model. Chan, Chen, and Lakonishok (2002) find in a sample of US domestic equity funds in 1976-1997 that fund styles measured on the basis of Fama-French (1993) three-factor model tend to cluster around a broad market benchmark. When deviating, funds are more likely to favor growth stocks with good recent performance. There is some consistency in styles, although funds with poor past performance are more likely to change styles. Using daily returns of US domestic equity funds in 1985-1995, Lynch and Musto (2000) find that the changes in the factor loadings from Carhart (1997) four-factor model are larger for funds in the bottom performance quartile than for the other funds. Poorly performing funds tend to increase investments in growth stocks, while good performers are likely to decrease their momentum loadings. The change in strategy as well as managerial replacement among the poor performers seem to lead to the performance improvement. Note, however, that the results of these studies should be taken with caution, since they are also a subject to the critique of Busse (2001) that statistical tests should account for the auto-correlation and cross-correlation in fund returns.

Several studies investigate the gaming behavior by mutual fund managers around the year-ends. Using the database with daily returns of US diversified equity funds in 1985-1997, Carhart et al. (2002) find strong evidence that some fund managers mark up their holdings at the last trading day of the year to improve a calendar-year performance (similar although weaker effects are also found at the quarter-ends). By trading aggressively at the end of the trading day, a manager pumps up the closing prices of his portfolio holdings, which determine the fund's net asset value and daily return. The authors show that funds with the greatest ability and the most incentive to improve their performance rankings are more active in marking up. Musto (1999) presents evidence of window dressing by managers of money market funds in 1987-1997. He demonstrates that funds allocating between government and private issues tend to increase their government holdings around the disclosure dates (at the fiscal year-end and six months later).

Since fund performance is reported on the net-of-fee basis, a manager can improve the fund's relative performance by waiving a part of his contracted fee. Christoffersen (2001) documents that over half of US money market funds waived fees in 1990-1995. This effect is economically significant: institutional funds waive almost half of their contracted advisory fees (19 basis points per year), while retail funds waive about two-thirds of their contracted fees (33 basis points per year). Fee waivers allow managers to flexibly react throughout the year to changes in relative performance, which affect fund flows. The link between fund performance and fee waivers appears to be especially strong and statistically significant among poorly performing funds, for which lower performance is associated with larger amounts of waived fees. A convex flow-performance relationship seems to encourage well-performing retail funds to increase waivers as a function of their performance. However, the fee waivers remain largely flat among well-performing institutional funds. This is interpreted as evidence of greater price competition among institutional funds than among retail funds.

Chapter 3

The dynamics of the impact of past performance on mutual fund flows[1]

3.1 Introduction

Many studies have recently analyzed the determinants of the behavior of mutual fund investors, concentrating on the relation between net inflows to mutual funds and their past performance. This research is of obvious relevance both for managers of mutual funds and their regulators. For the managers, it is important to know the factors that determine the total net assets under management which drive their compensation. The regulators should be aware of the incentives for risk-taking induced to managers by the existing investor behavior patterns.

The stylized findings indicate a clear positive impact of both risk-adjusted as well as raw past performance on subsequent net inflows (see, e.g., Ippolito, 1992, and Gruber, 1996). The relationship appears convex, indicating that most of the inflows are attracted by the best performing funds (see, e.g., Chevalier and Ellison, 1997, and Sirri and Tufano, 1998). Flows are also directly related to fund visibility, as funds belonging to larger families (see Sirri and Tufano, 1998) and funds advertising in the financial magazines (see Jain and Wu, 2000) tend to attract larger flows. Moreover, flows into a fund are found to be positively related to the performance of the fund family, measured, e.g., as average performance within the family (see, e.g., Ivkovic, 2000) or through the presence of star performers in the family (see, e.g., Nanda, Wang, and Zheng, 2000).

[1]Part of this research was carried out under a grant of the BSI Gamma Foundation, which is gratefully acknowledged.

Barber, Odean, and Zheng (2001) find that fund flows are more sensitive to the salient fees such as loads and commissions than to operating expenses. Del Guercio and Tkac (2002) document that mutual fund investors use less sophisticated measures of fund performance than pension fund clients.

The findings on the flow-performance relationship can be compared to the predictions based on the literature on performance persistence of mutual funds (see, e.g., Hendricks, Patel, and Zeckhauser, 1993, Wermers, 2000, Baks, Metrick, and Wachter, 2001, and many others). In general, these studies find strong evidence of persistence among bad performers and mixed evidence for consistent superior persistence. This implies that the relationship between fund flows and past performance should be the strongest among the worst-performing funds, which is opposite to the observed pattern (see Sirri and Tufano, 1998). This difference can be explained by a number of institutional and psychological factors, which prevent large outflows from funds with bad past performance. Market frictions such as the presence of search costs, back-end load charges, tax considerations, and restrictions of the investment retirement plans increase the transaction costs of withdrawing money from the poorly performing funds, while status-quo bias (see Zeckhauser, Patel, and Hendricks, 1991) and cognitive dissonance bias (see Goetzmann and Peles, 1997) make investors ignore information about bad fund performance.

Most studies referred to above focus on the impact of average past performance on fund flows at an annual frequency. In contrast, we analyze the full dynamic structure of the flow-performance relationship at the monthly frequency. As noted by Geweke (1978), low (e.g., annual) frequency analysis of the flow-performance relationship can be biased in a non-trivial way, if the true link is at higher frequency, and clearly cannot reveal the full high frequency lag structure. We find that performance from 6 to 8 months ago has the strongest impact on net flows to US growth funds. The performance during the most recent quarter appears less important than performance during the rest of the past year. This suggests that information dissemination takes time and some investors react to fund performance with a certain lag. The first three years of performance history account for about 90% of the total impact of past performance on flows. Moreover, almost all studies referred to above assume that the flow-performance sensitivity is constant. We find evidence that relative flows of small and, to a lesser extent, young funds are much more sensitive to past performance than larger and older funds.

As stated above, many studies (e.g., Chevalier and Ellison, 1997, and Sirri and Tufano, 1998) have found a convex flow-performance relationship, indicating that funds with top recent performance attract most of the inflows. This stylized finding, together with the fact that the manager's compensation is typically a percentage of the fund's net assets (see, e.g., Khorana, 1996), has lead to the hypothesis that managers of funds with poor performance in the first half of the year have an incentive to increase risk in the second half of the year. Clearly, this incentive is likely to be stronger for managers who find themselves at mid-year close to the point where the flow-performance sensitivity changes most. Tests of the above hypothesis are reported, e.g., by Brown, Harlow and Starks (1996), Busse (2001), as well as in Chapter 5 of this thesis. We find in the present chapter that the convexity of the flow-performance relationship is robust to allowing for more flexible dynamic lag structures and dependence of the flow-performance sensitivity on age and size of the fund. This convexity appears to be mostly due to the difference in flows between the top performing half of the funds and bottom performing half of the funds. However, within each of these two segments the flow-performance relationship is close to linear, which suggests that primarily funds with the average performance have incentives to take excessive risk.

Finally, we find that the return on systematic risk factors in the last two years or so has a small positive impact on flows in excess of the impact of the risk-adjusted returns. This might indicate that mutual investors are style timers.

The structure of the remainder of the chapter is as follows. Section 3.2 describes the data set and methodology and discusses the relation between the typical model used in the literature and our basic model specification. In Section 3.3, we compare the empirical results based on the two models. We also discuss our findings concerning the lag structure of the flow-performance relationship. In Section 3.4, we estimate the lag structure over a longer period (from 1976 to 1998), using quarterly data on funds' total net assets in 1976-1990. Sections 3.5 and 3.6 present the results concerning the convexity of the flow-performance relationship and additional impact of raw returns on flows (in excess of the risk-adjusted returns), respectively. Section 3.7 concludes.

3.2 Data and methodology

The data employed in our analysis are provided by Micropal. The data set includes the month of fund foundation, total net assets, and total returns of the US funds for the period January 1970 to December 1998. While returns are available at the monthly frequency throughout this period, total net asset values are available at monthly frequency from December 1990 and at quarterly frequency in 1970-1990. The main sample period in our study is consequently taken as January 1991 to December 1998. In Section 3.4, we also incorporate the quarterly data on funds' total net assets in a period before 1991. In order to avoid heterogeneity based on differences in fund styles, we perform the analysis on US growth funds only. Since we use a five-year horizon for fund performance, our analysis is restricted in each month to the funds with at least five years of the return history. Note that we have annualized monthly returns and flows in order to make our results comparable to existing evidence, which is based on the annual data.

In order to reduce the impact of typos and mergers, we exclude from our data set 1% of outliers based on net relative flows (the largest 0.5% of observations and the smallest 0.5% of observations).[2] In order to concentrate on the flow-performance relationship for moderately sized funds and avoid that the results are determined by outliers, we also exclude 1% of the observations with the largest size, which belong to only 6 funds and span from 11 to over 80 billion dollars. Table 3.1 presents descriptive statistics of the fund characteristics. During the main sample period (1991-1998), an average fund had \$732 million of assets and experienced an inflow of \$50 million or 5.4% per year, ranging from \$215 million outflow for the bottom quintile to \$460 million inflow for the top quintile.

Note that our data set contains only funds that were still in operation in the beginning of 1999 and is survivorship biased. However, it is straightforward to show that it does not affect the consistency of OLS or WLS estimates, if past flows do not influence the probability of fund survival in a joint regression with returns. This assumption is fully in line with the empirical findings in Brown and Goetzmann (1995). Not surprisingly, Chevalier and Ellison (1997), Goetzmann and Peles (1997), and Sirri and Tufano (1998) find the same results for survivorship biased and unbiased samples.

[2]Such approach of a truncated regression to curb the influence of flow outliers is also used in Edelen (1999).

Table 3.1: **Summary statistics of US growth funds**

The table reports summary statistics of US growth funds during the main sample period (1991-1998). Columns 2 and 3 report mean and standard deviation, while the last two columns present means of the fund characteristics in the respective top and bottom quintiles. Jensen's alpha and unsystematic risk (measured as standard deviation of the residuals) are calculated on the basis of the the four-factor model of Carhart (1997). Note that Jensen's alpha, nonsystematic risk, absolute and relative flows are annualized.

Fund characteristic	Mean	Std. Dev.	Mean (bottom quintile)	Mean (top quintile)
Absolute flow, \$mln	49.87	382.57	-214.95	460.74
Relative flow, %	5.36	40.59	-36.00	61.34
Total Net Assets, \$mln	732.35	1228.44	18.25	2617.05
Age, years	16.27	10.37	5.98	31.29
Jensen's alpha, %	-0.13	3.24	-4.50	4.42
Nonsystematic risk, %	17.53	6.46	10.56	27.53

Traditionally (see, e.g., Gruber, 1996), net absolute flows are defined as the change in fund assets net of reinvested dividends:

$$F_{i,t} = TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t}), \tag{3.1}$$

where $TNA_{i,t}$ denotes fund i's total net assets at the end of month t and $R_{i,t}$ is return of fund i in month t. Similarly, net relative flows are defined as a net percentage growth of fund assets:

$$f_{i,t} = \frac{TNA_{i,t} - (1 + R_{i,t})TNA_{i,t-1}}{TNA_{i,t-1}} = \frac{F_{i,t}}{TNA_{i,t-1}}. \tag{3.2}$$

Both definitions are based on an assumption that all investor earnings are automatically reinvested in the fund and flows occur at the end of month t. Due to the low autocorrelation in monthly returns, flows occuring at other instances during the month will not bias any of our results. To account for the impact of the inflation, we deflate funds' total net asset values by the US consumer price index and convert them into equivalent US dollars, as of December 1990 before computing flow measures.

Almost all studies referred to in the previous section analyze flows at the annual frequency, assuming that the sensitivity of flows with respect to past performance is the same for all funds. Thus, the standard model in the literature specifies net relative flows as a linear function of past performance and a set of control variables:

$$f_{i,t} = a + b_1 r_{i,t-1} + ... + b_K r_{i,t-K} + c' x_{i,t-1} + u_{i,t}, \tag{3.3}$$

where $r_{i,t-i}$ is some measure of fund i's performance (e.g., raw return, Jensen's alpha, or corresponding ranking) in period $t-i$ and $x_{i,t-1}$ includes such variables as fund size, age, fees, a measure of riskiness, and performance of other funds in the family. The assumption that the flow-performance sensitivity coefficients b_1 to b_K do not depend on fund characteristics, such as size and age, is clearly restrictive. Moreover, one should keep in mind that small funds have extreme relative flows that dominate OLS estimates. Unless heteroskedasticity-consistent standard errors are computed, inference based on OLS estimates will be biased. For efficiency reasons, we model the variance of the error term and compute weighted least squares estimates.

In this chapter, we try to model the impact of past performance on flows in a less restrictive way. We write our model first in terms of absolute flows, where we specify both the performance-unrelated part and the flow-performance sensitivity of the flow model as polynomials in logs of fund size and age:

$$F_{i,t} = G(TNA_{i,t-1}, age_{i,t-1}) + H(TNA_{i,t-1}, age_{i,t-1}) \sum_{j=1}^{60} w_j RAR_{i,t-j} + e_{i,t}, \tag{3.4}$$

where functions G and H approximate the unknown functional form of the performance-unrelated and performance-related parts of the relationship between flows and performance. The empirical results to be presented later suggest that a specification with second-order polynomials in logs of fund size and age suffices.[3]

Equivalently, we can rewrite the model in terms of relative flows. After dividing both sides of (3.4) by $TNA_{i,t-1}$, we obtain

$$f_{i,t} = g(TNA_{i,t-1}, age_{i,t-1}) + h(TNA_{i,t-1}, age_{i,t-1}) \sum_{j=1}^{60} w_j RAR_{i,t-j} + \tilde{e}_{i,t}, \tag{3.5}$$

where $g(\cdot) \equiv G(\cdot)/TNA_{i,t-1}$, $h(\cdot) \equiv H(\cdot)/TNA_{i,t-1}$, and $\tilde{e}_{i,t} \equiv e_{i,t}/TNA_{i,t-1}$.

Fund performance over the past five years is measured as a weighted sum of past risk-adjusted returns defined on the basis of the four-factor model with the market, size,

[3]The joint hypothesis that the third-order terms are zero is not rejected at the conventional confidence level.

book-to-market, and one-year momentum factors[4], as in Carhart (1997):

$$RAR_{i,t-j} \equiv R_{i,t-j} - R^f_{t-j} - \sum_{l=1}^{4} \hat{\beta}^l_i F^l_{t-j}, \tag{3.6}$$

where $F_t = (R^m_{t-j} - R^f_{t-j}, SMB_{t-j}, HML_{t-j}, MOM_{t-j})$ and $\hat{\beta}^1_i, ..., \hat{\beta}^4_i$ are estimated using all observations available for a given fund. In order to ensure the smoothness of the impulse response function, we impose a polynomial structure on the performance coefficients. We approximate the distribution of the lag coefficients on risk-adjusted returns by a polynomial of the p-th order:

$$w_j = \sum_{k=0}^{p} \theta_k k! j^{-k} \text{ for } j = 1, ..., 60. \tag{3.7}$$

The empirical results indicate that $p = 5$ suffices. Since we expect that the impact of past performance disappears after at most five years, we impose the end-point restriction that $w_{61} = 0$. In order to identify the model, we normalize the weights, so that the average of the performance coefficients is equal to one: $\frac{1}{60}\sum_{j=1}^{60} w_j = 1$. The performance coefficients represent the weights with which investors take past performance into account. If all weights are equal to each other (i.e. $\theta_k = 0$ for $k > 0$), the weighted sum of risk-adjusted returns in (3.4) equals Jensen's alpha over a five-year estimation period.

Throughout the chapter, we compute weighted least squares estimates where the variance of $e_{i,t}$ is modelled as

$$Var(e_{i,t}) = \exp U(TNA_{i,t-1}, age_{i,t-1}), \tag{3.8}$$

with U being a second-order polynomial in logs of fund size and age. This specification reflects that the disturbances of both the absolute flow specification (3.4) and the relative flow specifications (3.5) are heteroskedastic, in contrast to what is often assumed in the literature. The coefficients of the function U are estimated on the basis of the OLS residuals.

We estimate the model parameters in (3.4) by means of a concentrated least-squares approach. For the pre-specified values of the parameters in the function H, the model (3.4) is linear in the remaining parameters. Therefore, the least squares estimates can

[4]We thank Kenneth R. French for the opportunity to use the factor returns provided at his website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

Table 3.2: **Flow-performance relationship: a typical model in the literature**
The table reports the estimated coefficients and standard errors (in the parentheses) based on the typical model used in the literature (3.9) for the period 1991-1998. The dependent variable is fund net relative flow. The independent variables include a constant, log of fund size, log of fund age, and five-year Jensen's alpha. Note that relative flows and Jensen's alpha are annualized.

$Const$	15.52	(1.26)
$\log TNA_{i,t-1}$	-0.64	(0.15)
$\log age_{i,t-1}$	-4.20	(0.41)
$\alpha_{i,t}$	3.79	(0.08)

conveniently be computed by numerically maximizing the concentrated sum of squares over the parameters in the function H.

3.3 Basic results

A typical example of a specification considered in the literature is

$$f_{i,t} = a + c_1 \log TNA_{i,t-1} + c_2 \log age_{i,t-1} + b\alpha_{i,t} + u_{i,t}, \tag{3.9}$$

where fund performance is measured as Jensen's alpha $\alpha_{i,t}$ over a five-year period. This is equivalent to imposing $w_1 = ... = w_{60} = \frac{1}{60}$ in our basic specification (3.5). Moreover, the performance-unrelated part of the model, g in (3.5), is specified as being linear in logs of fund size and age, while the flow-performance sensitivity, h in (3.5), is simply taken to be constant.

The estimation results for the model (3.9) are reported in Table 3.2. In line with the existing evidence, we find that better performing funds, smaller funds, and younger funds attract larger relative flows. The dependence of the performance-unrelated part of the specification for relative flows (the function g) on fund size and age is illustrated graphically in Panel A of Figure 3.1.[5] This figure shows that the expected flows of funds with a neutral past performance (i.e., with Jensen's alpha equal to zero) range

[5]In all graphs depicting expected fund flows as a function of size and age, size and age axes start from \$250 million and 5 years, respectively. We exclude the segment of the smallest funds because of the large standard errors of their expected relative flows.

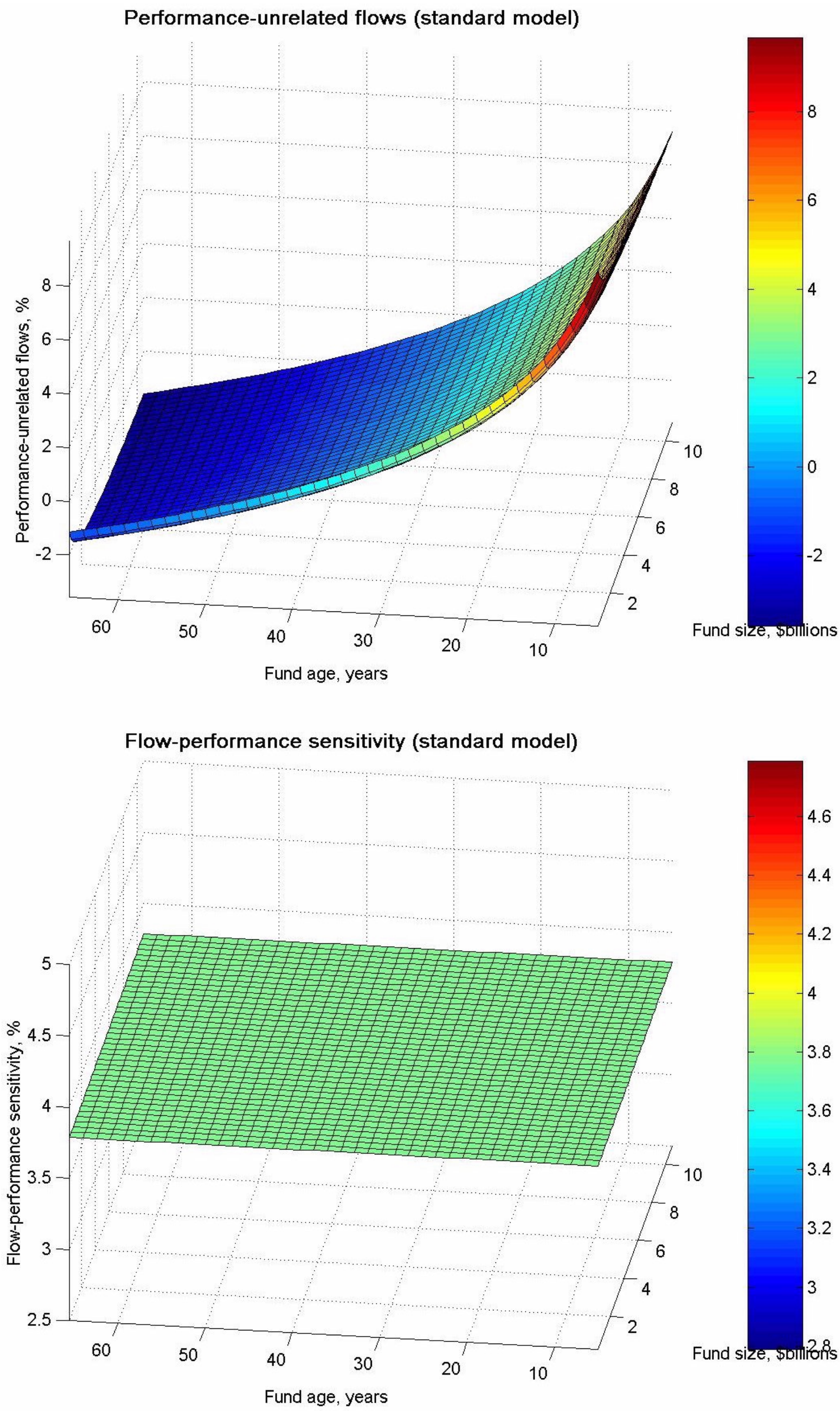

Figure 3.1: **Expected fund flows as a function of size and age (standard model)** Panels A and B show the performance-unrelated flows (flows of a fund with zero Jensen's alpha) and flow-performance sensitivity (change in fund flows due to 1% increase in Jensen's alpha), predicted by the standard model in the literature (3.9). The performance-unrelated flows are modelled as linear in logs of size and age. The flow-performance sensitivity is assumed to be constant.

from -3% for large old funds to 9% for small young funds. The model imposes that the flow-performance sensitivity of relative flows (the function h) is constant over funds (see Panel B in Figure 3.1). It predicts that 1% change in Jensen's alpha will lead to 3.8% change in expected relative flows, which is consistent with the findings in, e.g., Chevalier and Ellison (1997). It can readily be seen that the current specification may be too restrictive, since it predicts negative flows for old funds with neutral performance and the same sensitivity for the smallest and the largest funds.

The estimation results for the basic specification as put forward in (3.4) to (3.8) are presented in Table 3.3. All coefficients are highly significant, which allows us to reject the hypothesis that the sensitivity of flows to performance is the same for all funds. The expected performance-unrelated flows rise from 0.5% for large old funds to 2% for large young funds and 10-16% for small young funds (see Panel A in Figure 3.2). This pattern looks much more reasonable than the one based on the restrictive specification (3.9) typically used in the literature. The peak of 16% in the segment of smallest youngest funds can be explained by the observation that smallest funds tend to have negative Jensen's alpha in a range from -2% to -3%, which compensates the peak. The flow-performance sensitivity is also higher for smaller and younger funds (see Panel B in Figure 3.2). It ranges from 0.5% for large old funds to 5-8% for small young funds. Thus, flows to small and, to some extent, young funds appear much more sensitive to past performance than flows to large and old funds. One possible explanation is that investors invest approximately equal dollar amounts in the best performing funds, irrespective of their current size, which would make the relative flow to small funds much more sensitive to past performance. Moreover, investors may be more sensitive to the recent performance of young funds, since they have not yet obtained the reputation established by the old funds.

The hypothesis that average past risk-adjusted performance over a five-year period determines subsequent inflows, i.e., that $\theta_k = 0$ for $k > 0$, is strongly rejected. The impact of past performance on subsequent flows that is implied by the estimated θ's is illustrated in Panel A of Figure 3.3. The information content of past performance rises during the first eight months and then gradually decreases towards zero. As indicated by the confidence bands in Panel A of Figure 3.3, the specification on the basis of average past risk-adjusted performance at annual or quarterly frequency is strongly rejected. Current flows are most strongly affected by the performance from 6 to 8 months ago. The

Table 3.3: **Lag structure of the flow-performance relationship**

The table reports the estimated coefficients and standard errors (in the parentheses) based on the specification (3.4) for the period 1991-1998 (see columns 2 to 5 in Panel A and columns 2 to 3 in Panel B) and on the specifications (3.4) and (3.11) for the period 1976-1998, including the period 1976-1990 with quarterly data on flows (see the last four columns in Panel A and the last two columns in Panel B). In both specifications, the dependent variable is fund net absolute flow. The independent variables include the performance-unrelated term and flow-performance sensitivity times weighted sum of past 60 monthly risk-adjusted returns. Both the performance-unrelated term and flow-performance sensitivity are modelled as a quadratic function of logs of fund size and age (G-function and H-function, respectively). The performance coefficients are restricted to lie on a polynomial of the fifth order (see (3.7)). Note that flows and Jensen's alpha are annualized.

Panel A. Size and age coefficients

	1991-1998				1976-1998			
	G-function		H-function		G-function		H-function	
$Const$	202.42	(14.69)	66.15	(5.15)	198.49	(12.74)	39.5	(3.93)
$\log TNA$	61.96	(4.35)	25.53	(1.04)	57.23	(3.96)	12.56	(1.11)
$\log^2 TNA$	4.93	(0.43)	2.53	(0.08)	4.23	(0.38)	1.19	(0.10)
$\log age$	-77.63	(9.01)	-13.89	(3.09)	-79.58	(6.60)	-11.8	(1.66)
$\log^2 age$	7.73	(1.48)	1.00	(0.48)	7.69	(1.06)	1.41	(0.23)
$\log TNA * \log age$	-11.07	(1.10)	-2.22	(0.26)	-11.13	(0.88)	-1.12	(0.18)

Panel B. Performance coefficients

	1991-1998		1976-1998	
θ_0	-0.02	(0.00)	-0.02	(0.00)
θ_1	1.19	(0.05)	1.17	(0.08)
θ_2	-4.40	(0.41)	-4.25	(0.69)
θ_3	4.62	(0.66)	4.42	(1.13)
θ_4	-1.58	(0.29)	-1.51	(0.51)
θ_5	0.15	(0.03)	0.14	(0.06)

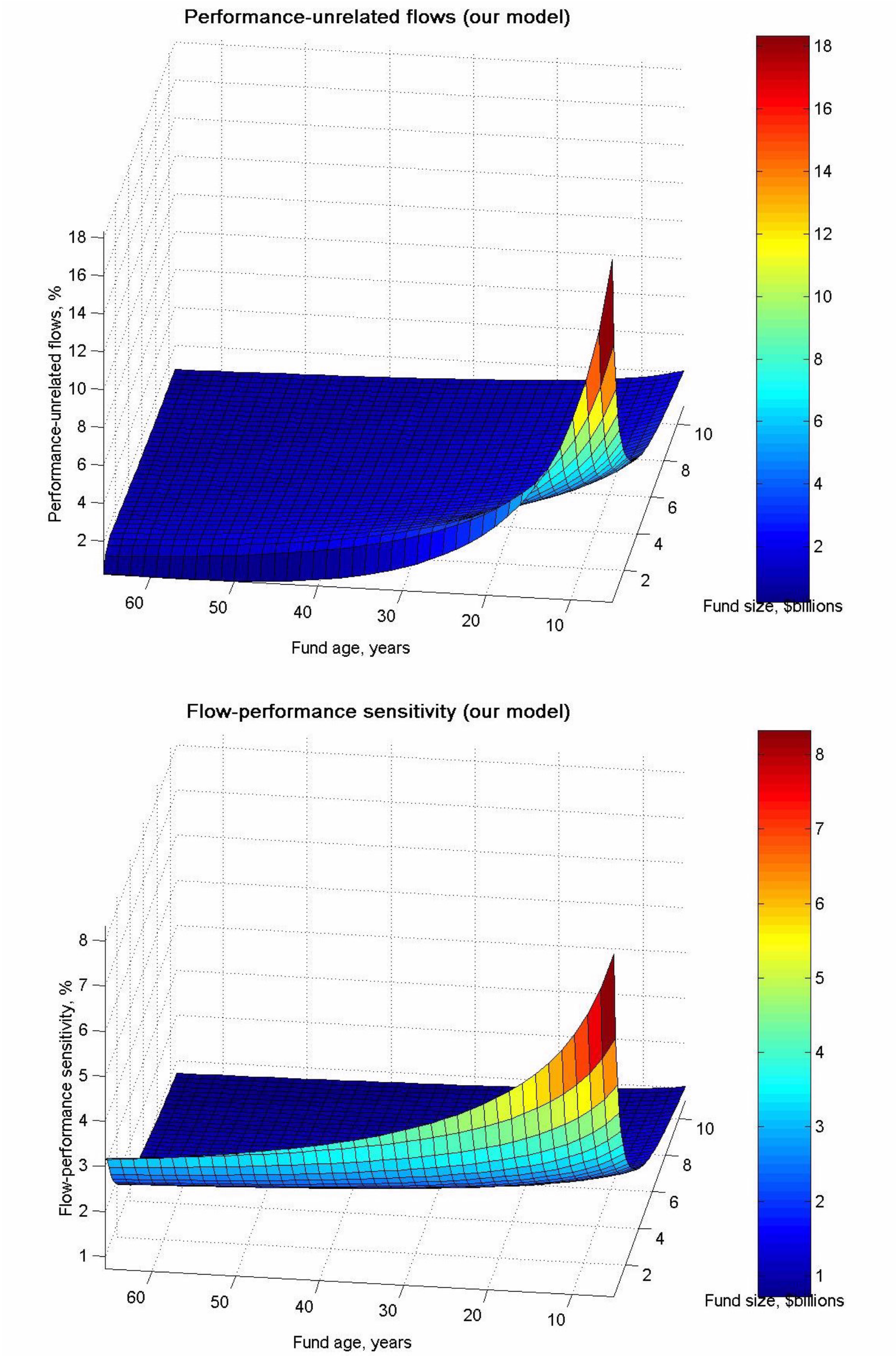

Figure 3.2: **Expected fund flows as a function of size and age (our model)**
Panels A and B show the performance-unrelated flows (flows of a fund with zero Jensen's alpha) and flow-performance sensitivity (change in fund flows due to 1% increase in Jensen's alpha), based on the basic specification (3.4) for the period 1991-1998. Both the performance-unrelated flows and flow-performance sensitivity are modelled as a quadratic function in logs of fund size and age.

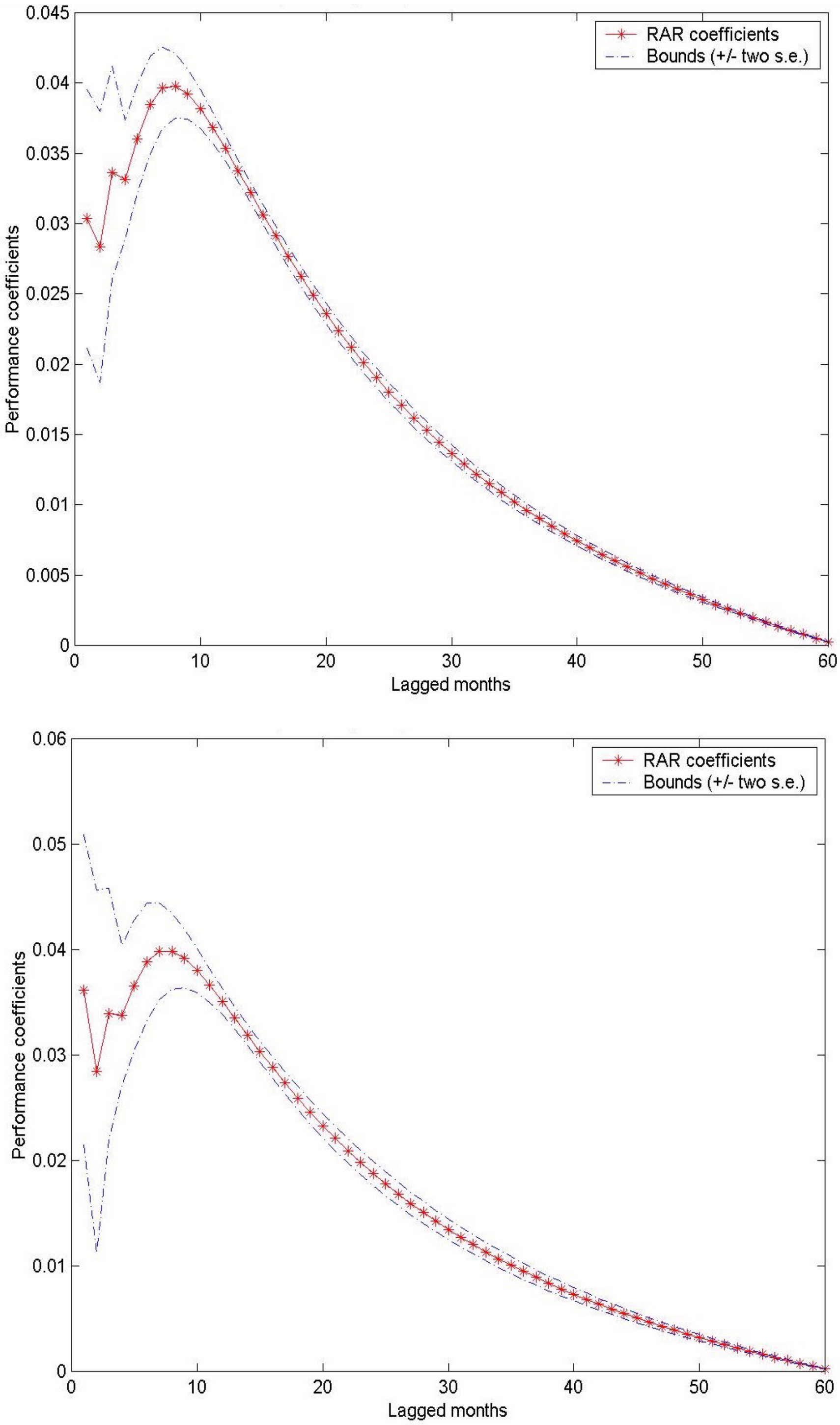

Figure 3.3: **Impact of past performance on flows**

The graph shows the lag structure of the impact of past 60 monthly risk-adjusted returns on current flows. The performance coefficients are restricted to lie on a polynomial of the fifth order (see (3.7)). Panels A and B are based on the basic specification (3.4) for the period 1991-1998 and modified specification (3.11) for the period 1976-1998 (including the period 1976-1990 with quarterly data on flows), respectively.

performance during the most recent quarter appears less important than performance during the rest of the past year. This suggests that information dissemination takes time and some investors react to fund performance with a certain lag. Consistent with previous findings (see, e.g., Sirri and Tufano, 1998), fund performance during the most recent year has the strongest impact on flows, accounting for 43% of the total impact. The sensitivity of flows to past performance fades away after a period of three years.

Formal tests of statistical significance of the results are reported in Panel A of Table 3.4. In the lower diagonal part of the table we present p-values of Wald tests of the hypothesis that the impact of past performance i months ago equals that of j months ago: $H_0 : w_i = w_j$. The upper diagonal part of the table reports p-values of Wald tests of the hypothesis that the impact of average past performance i quarters ago equals that of j quarters ago: $\mathrm{H}_0 : w_{3i-2} + w_{3i-1} + w_{3i} = w_{3j-2} + w_{3j-1} + w_{3j}$. The table indicates that the impact of average performance three quarters ago differs significantly from the impact of average performance one, two and four quarters ago (p-values of 0.0058, 0.0019, and 0.0001, respectively).

3.4 Extension of the sample period using quarterly data on Total Net Assets

In this section, we estimate the lag structure of the flow-performance relationship over a longer time span, from 1976 to 1998. Since the data on funds' total net assets before 1991 are only available at quarterly frequency, we calculate quarterly flows from the monthly flows over that quarter. After adding the basic equation (3.4) for three consecutive months and dividing by 3, we obtain

$$\begin{aligned}&\tfrac{1}{3}(F_{i,t} + F_{i,t-1} + F_{i,t-2}) = \tfrac{1}{3}\textstyle\sum_{l=1}^{3} G(TNA_{i,t-l}, age_{i,t-l}) + \\ &\tfrac{1}{3}\textstyle\sum_{l=1}^{3}[H(TNA_{i,t-l}, age_{i,t-l})\sum_{j=1}^{60} w_j RAR_{i,t-j-l+1}] + \tfrac{1}{3}\sum_{l=1}^{3} e_{i,t-l+1}.\end{aligned} \tag{3.10}$$

Since monthly flows in (3.4) are annualized, the left-hand side of the equation is equal to the annualized quarterly net flow realized during months $t-2$ to t: $F_{i,t:t-2}$. Since monthly net assets values are not observed for the first part of the sample period we approximate the specification by

$$\begin{aligned}&F_{i,t:t-2} = G(TNA_{i,t-3}, age_{i,t-3}) + \\ &H(TNA_{i,t-3}, age_{i,t-3})\textstyle\sum_{j=1}^{60} w_j(RAR_{i,t-j} + RAR_{i,t-j-1} + RAR_{i,t-j-2}) + \tilde{e}_{i,t}.\end{aligned} \tag{3.11}$$

Table 3.4: **Tests of the hypotheses about the lag structure of the flow-performance relationship**

Panels A and B of the table describe tests based on the specification (3.4) for the period 1991-1998 and on the specifications (3.4) and (3.11) for the period 1976-1998 (including the period 1976-1990 with quarterly data on flows), respectively. The lower diagonal part of the table presents p-values of the tests of the hypothesis that the impact of past performance i months ago equals that of j months ago, $H_0 : w_i = w_j$. The upper diagonal part of the table reports p-values of the tests of the hypothesis that the impact of past performance i quarters ago equals that of j quarters ago, $H_0 : w_{3i-2} + w_{3i-1} + w_{3i} = w_{3j-2} + w_{3j-1} + w_{3j}$.

Panel A. Sample period 1991-1998

$i\backslash j$	1	2	3	4	5	6	7	8	9	10
1	-	0.1144	0.0058	0.0407	0.6130	0.2534	0.0075	0.0001	0.0000	0.0000
2	0.7594	-	0.0019	0.5891	0.0554	0.0001	0.0000	0.0000	0.0000	0.0000
3	0.6053	0.3921	-	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4	0.5908	0.4109	0.8621	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5	0.2478	0.1773	0.5767	0.0237	-	0.0000	0.0000	0.0000	0.0000	0.0000
6	0.0933	0.0684	0.2952	0.0023	0.0000	-	0.0000	0.0000	0.0000	0.0000
7	0.0515	0.0360	0.1928	0.0008	0.0001	0.0029	-	0.0000	0.0000	0.0000
8	0.0465	0.0297	0.1743	0.0011	0.0014	0.0692	0.6831	-	0.0000	0.0000
9	0.0613	0.0358	0.2083	0.0041	0.0238	0.4620	0.4817	0.0446	-	0.0000
10	0.0993	0.0549	0.2967	0.0211	0.1832	0.8010	0.0851	0.0020	0.0000	-

Panel B. Sample period 1976-1998

$i\backslash j$	1	2	3	4	5	6	7	8	9	10
1	-	0.4337	0.115	0.3386	0.8073	0.1388	0.0091	0.0003	0.0000	0.0000
2	0.5319	-	0.0804	0.9441	0.1407	0.0053	0.0001	0.0000	0.0000	0.0000
3	0.8241	0.6345	-	0.0106	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4	0.7603	0.6036	0.9747	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5	0.9561	0.4003	0.6952	0.1710	-	0.0000	0.0000	0.0000	0.0000	0.0000
6	0.7221	0.2653	0.5026	0.0689	0.0102	-	0.0000	0.0000	0.0000	0.0000
7	0.6208	0.2084	0.4179	0.0462	0.0205	0.1085	-	0.0000	0.0000	0.0000
8	0.6200	0.2000	0.4067	0.0563	0.0768	0.3881	0.9953	-	0.0000	0.0000
9	0.6883	0.2237	0.4511	0.1014	0.2472	0.8482	0.4972	0.1449	-	0.0000
10	0.8047	0.2738	0.5440	0.2081	0.5727	0.6794	0.1966	0.0358	0.0041	-

We estimate the flow-performance relationship combining the modified model (3.11) for quarterly flows during the first part of the sample period (1976-1990) and basic model (3.4) for monthly flows during the second part of the sample period (1991-1998). We impose the same identifying and end-point restrictions and polynomial structure, as before (see (3.6) and (3.7)). Table 3.3 and Panel B in Table 3.4 report the results, which are very similar to those based only on monthly flows in 1991-1998. The parameters in the G and H functions are estimated more precisely due to the additional information about the quarterly flows in 70s and 80s. The lag structure has the same general shape as before, with the impact of performance on flows rising during the first three quarters ago and fading subsequently to zero (see Panel B in Figure 3.3). Note that the standard errors of the performance coefficients have somewhat increased, which could be due to temporal changes in the lag structure.

3.5 Convexity of the flow-performance relationship

In the previous sections, we assumed that the flow-performance relationship is the same for good and bad performers. In this section, we reexamine the existing evidence on nonlinearity of the flow-performance relationship by allowing the impact of past performance to be different in each of five segments corresponding to performance quintiles based on five-year Jensen's alpha. The kink points between the segments are the quintile points of the estimated distribution of Jensen's alphas. Formally, we rewrite the basic model (3.4) as

$$F_{i,t} = G(TNA_{i,t-1}, age_{i,t-1}) + H(TNA_{i,t-1}, age_{i,t-1}) \sum_{p=1}^{5} \sum_{j=1}^{60} w_j(p) RAR_{i,t-j} + e_{i,t}, \quad (3.12)$$

where we allow each performance coefficient w_j to be different across quintiles. We impose the identifying restriction $\frac{1}{300} \sum_{p=1}^{5} \sum_{j=1}^{60} w_j(p) = 1$ and specify a polynomial structure of the fifth order for the performance coefficients in each segment:

$$w_j(p) = \sum_{k=0}^{5} k! \theta_k(p) j^{-k} \text{ for } j = 1, ..., 60. \quad (3.13)$$

The end-point restriction is imposed in every segment.

Table 3.5 presents the results. The impulse response function is very similar in all quintiles, as the flow-performance sensitivity peaks in a period 6-8 months ago and

Table 3.5: **Quintile-specific lag structure of the flow-performance relationship** The table reports the estimated coefficients and standard errors (in the parentheses) based on the model (3.12) for the period 1991-1998. The dependent variable is fund net absolute flow. The independent variables include the performance-unrelated term and flow-performance sensitivity times weighted sum of past 60 monthly risk-adjusted returns. Both the performance-unrelated term and flow-performance sensitivity are modelled as a quadratic function of logs of fund size and age (G-function and H-function, respectively). The quintile-specific performance coefficients are restricted to lie on a polynomial of the fifth order (see (3.13)). The quintiles are defined on the basis of the five-year Jensen's alpha. Note that flows and Jensen's alpha are annualized.

Panel A. Size and age coefficients

	G-function		H-function	
$Const$	165.42	(20.75)	62.94	(5.37)
$\log TNA$	48.48	(8.33)	24.4	(1.12)
$\log^2 TNA$	3.69	(0.87)	2.44	(0.10)
$\log age$	-67.59	(7.86)	-13.45	(3.32)
$\log^2 age$	6.91	(1.33)	1.01	(0.54)
$\log TNA * \log age$	-9.45	(1.07)	-2.16	(0.28)

Panel B. Performance coefficients

Quintiles	Bottom		2nd		3rd		4th		Top	
θ_0	-0.01	(0.00)	-0.01	(0.00)	-0.02	(0.00)	-0.02	(0.00)	-0.02	(0.00)
θ_1	0.94	(0.17)	0.38	(0.19)	1.34	(0.23)	1.50	(0.18)	1.67	(0.14)
θ_2	-3.55	(0.95)	0.05	(1.22)	-4.82	(1.40)	-5.99	(1.09)	-6.74	(0.99)
θ_3	3.88	(1.36)	-1.37	(1.80)	4.71	(1.99)	6.62	(1.62)	7.76	(1.53)
θ_4	-1.39	(0.57)	0.81	(0.76)	-1.48	(0.82)	-2.33	(0.69)	-2.87	(0.67)
θ_5	0.13	(0.06)	-0.10	(0.08)	0.13	(0.09)	0.22	(0.07)	0.29	(0.07)

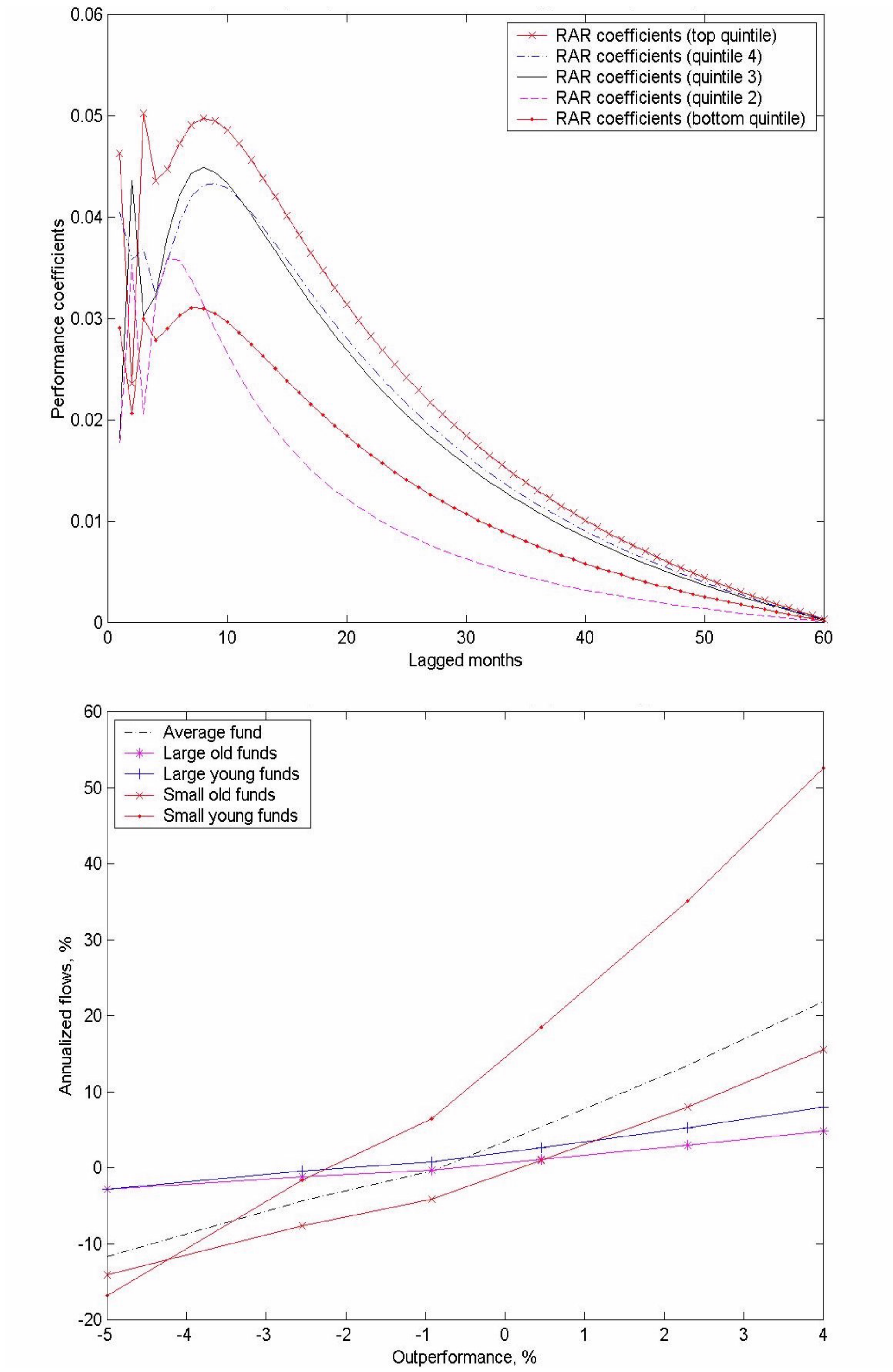

Figure 3.4: **Quintile-specific impact of past performance on flows**

Panel A shows the lag structure of the impact of past 60 monthly risk-adjusted returns on current flows, based on the model (3.12) for the period 1991-1998. The quintile-specific performance coefficients are restricted to lie on a polynomial of the fifth order (see (3.13)). The performance quintiles are defined on the basis of the five-year Jensen's alpha. Panel B depicts expected flows for funds with five different combinations of size and age. An average fund has an age of 16 years and $732 mln in assets. Small and large funds have a size of $250 mln and $8 bln, while young and old funds have an age of 5 and 50 years, respectively.

Table 3.6: **Tests of the hypotheses about the quintile-specific lag structure of the flow-performance relationship**

The table describes tests based on the specification (3.12) for the period 1991-1998. The lower diagonal part of the table presents p-values of the tests of the hypothesis that the average impact of past performance on flows of funds in quintile p equals that in quintile r, $H_0 : \sum_j w_j(p) = \sum_j w_j(r), \forall p, r$.

$p \backslash r$	1	2	3	4
1	-			
2	0.2795	-		
3	0.0327	0.0005	-	
4	0.0079	0.0000	0.7784	-
5	0.0000	0.0000	0.0877	0.1688

then converges towards zero (see Panel A in Figure 3.4). However, flows to the better performing funds appear much more sensitive to past performance than flows to badly performing funds. The hypothesis of linearity of the flow-performance relationship, formulated in terms of the average quintile-specific performance coefficients as $H_0 : \sum_j w_j(p) = \sum_j w_j(r)$ $(\forall p, r)$ is clearly rejected. The corresponding Wald test has a p-value below 0.0001. Thus, we find that the well-documented convexity of flows with respect to past performance found in other studies (see, e.g., Chevalier and Ellison, 1997, and Sirri and Tufano, 1998) is robust to allowing for dependence of this relationship on size and age of the fund. Apparently, this convexity is mostly due to the difference in flow-performance sensitivity between the top three and bottom two performance quintiles. As reported in the lower diagonal part of Table 3.6, all but one pair-wise differences in the average performance coefficients between the quintiles from these segments are significant at 1% level. The convexity pattern is illustrated in Panel B of Figure 3.4, which depicts relative flows as a function of the outperformance with respect to the market for an average fund and funds with different combinations of size and age. According to our model, an average fund is expected to lose about 12% in outflows when underperforming the market by 5% per year and is expected to gain about 18% in inflows when outperforming the market by 3% per year. Given the same performance, a small old fund would lose about 14% in outflows, while small young fund would attract about 40% inflows. As we saw before, flows to large funds are much less sensitive to past performance.

3.6 Impact of benchmark risk

So far, we have demonstrated a strong positive relation between fund flows and risk-adjusted returns. Similar results can be obtained for raw returns, since raw and risk-adjusted returns of US growth funds are highly correlated. An interesting question is whether raw returns add something to risk-adjusted returns in explaining fund flows. To answer that question, we add raw returns as one more performance measure to our basic model (3.4):

$$F_{i,t} = G(TNA_{i,t-1}, age_{i,t-1}) + H(TNA_{i,t-1}, age_{i,t-1})[\sum_{j=1}^{60} w_j RAR_{i,t-j} + \sum_{j=1}^{60} v_j R_{i,t-j}] + e_{i,t}, \tag{3.14}$$

where as before we impose a polynomial structure, the end-point restrictions, and an identifying restriction that $\frac{1}{60}\sum_{j=1}^{60}(w_j + v_j) = 1$. Note that since raw returns can be disentangled into the risk-adjusted and systematic risk components (factors are defined as in (3.6)):

$$R_{i,t-j} = RAR_{i,t-j} + R^f_{t-j} + \sum_{l=1}^{4} \hat{\beta}^l_i F^l_{t-j}, \tag{3.15}$$

we can rewrite the model (3.14) as

$$F_{i,t} = \tilde{G}(TNA_{i,t-1}, age_{i,t-1}) + H(TNA_{i,t-1}, age_{i,t-1})[\sum_{j=1}^{60} \tilde{w}_j RAR_{i,t-j} + \sum_{j=1}^{60} \tilde{v}_j (\sum_{l=1}^{4} \hat{\beta}^l_i F^l_{t-j})] + e_{i,t}, \tag{3.16}$$

where $\tilde{w}_j = w_j + v_j$, $\tilde{v}_j = v_j$, and $\tilde{G}(\cdot) = G(\cdot) + H(\cdot)\sum_{j=1}^{60} v_j R^f_{t-j}$.

The estimation results are presented in Table 3.7 and Figure 3.5. We find that both types of performance are positively related to flows, which is consistent with the findings of, e.g., Gruber (1996) and Sirri and Tufano (1998) based on annual data. The results indicate that the $\tilde{v}_j$ coefficients are small, but statistically significant. The recent outperformance on the systematic risk factors does yield additional inflow, which indicates that some mutual fund investors are style timers. The inclusion of the $\tilde{v}_j$ coefficients hardly affects the estimates of the $\tilde{w}_j$ coefficients.

Table 3.7: **Lag structure of the flow-performance relationship: raw vs. risk-adjusted performance**

The table reports the estimated coefficients and standard errors (in the parentheses) based on the model (3.14) for the period 1991-1998. The dependent variable is fund net absolute flow. The independent variables include the performance-unrelated term and flow-performance sensitivity times weighted sum of past 60 monthly risk-adjusted returns and past 60 raw returns. Both the performance-unrelated term and flow-performance sensitivity are modelled as a quadratic function of logs of fund size and age. The coefficients on raw and risk-adjusted returns are restricted to lie on a polynomial of the fifth order (see (3.7)). Note that flows and Jensen's alpha are annualized.

Panel A. Size and age coefficients

	G-function		H-function	
$Const$	11.90	(37.86)	68.25	(5.14)
$\log TNA$	-6.93	(13.69)	27.11	(1.03)
$\log^2 TNA$	-1.64	(1.32)	2.69	(0.09)
$\log age$	-30.19	(10.76)	-12.82	(3.12)
$\log^2 age$	3.70	(1.32)	0.69	(0.50)
$\log TNA * \log age$	-4.28	(1.44)	-2.29	(0.25)

Panel B. Performance coefficients

Returns	Risk-adjusted		Raw	
θ_0	-0.02	(0.00)	0.00	(0.00)
θ_1	1.18	(0.10)	0.08	(0.10)
θ_2	-4.94	(0.72)	0.09	(0.70)
θ_3	5.72	(1.09)	-0.45	(1.02)
θ_4	-2.10	(0.46)	0.25	(0.42)
θ_5	0.21	(-0.05)	-0.03	(0.04)

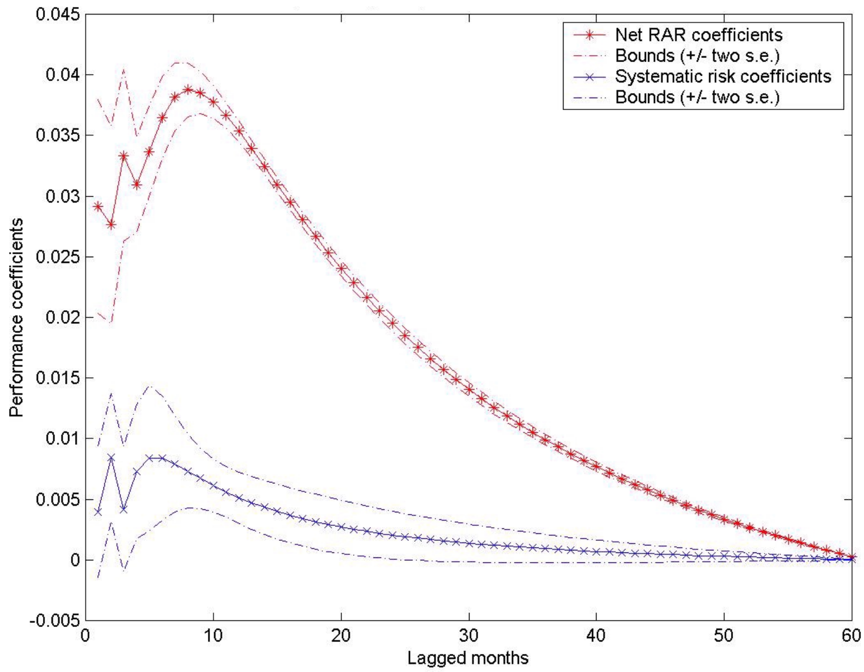

Figure 3.5: **Impact of raw and risk-adjusted performance on flows**

The graph shows the lag structure of the net impact of past 60 monthly risk-adjusted and raw returns on current flows, based on the model (3.14) for the period 1991-1998. The coefficients on raw and risk-adjusted returns are restricted to lie on a polynomial of the fifth order (see (3.7)).

3.7 Conclusion

In this chapter, we analyze the dynamic structure of the impact of past performance on fund flows. The flow-performance relationship is estimated at the monthly frequency, allowing for dependence of the sensitivity of flows to past performance on size and age of the fund. Traditional model specifications in the literature based on average past performance at annual or quarterly frequency are strongly rejected. We find that the impact of past performance on flows does not monotonically decay with time. Performance from 6 to 8 months ago seems to have the strongest impact on net flows to US growth funds. We observe that fund flows are less sensitive to performance during the most recent quarter than to performance during the remaining three quarters of the first year. This can be explained by information dissemination taking time and some investors reacting to fund performance with a certain lag. The impact of past performance on flows is mostly limited to the three most recent years of performance history, which accounts for about 90% of the total impact.

The well-documented convexity of the flow-performance relationship is robust to allowing for our more flexible dynamic lag structure and dependence of this relationship on size and age of the fund. This convexity seems to be mostly due to the difference in flows between the top performing half of the funds and bottom performing half of the funds. Within each of these two segments the flow-performance relationship appears close to linear, which suggests that funds with the average performance have more incentives to take excessive risk as a result of the convexity in the flow-performance relationship.

Finally, we find that performance on systematic risk factors has a small positive impact on flows in excess of the impact of the risk-adjusted returns. This suggests that some mutual fund investors are style timers choosing funds on the basis of their raw rather than risk-adjusted performance. However, this finding could be specific for our sample period (1991-1998), during most of which the systematic factors realized positive returns.

Chapter 4

The relative impact of different classification schemes on mutual fund flows

4.1 Introduction

The US mutual fund industry experienced tremendous growth during the past two decades. In 2001, there were 8307 mutual funds controlling over 6.9 trillion dollars in assets, which by far exceeds 665 funds with 241.4 billion dollars in assets in 1981 (Investment Company Institute, 2002). With so many funds around, investors face a difficult task of selecting a fund with the desired risk and performance profile. Mutual fund categories composed of funds with a similar investment approach help investors to simplify their decision problem. Investors often first choose the category that suits their preferences and then select the best fund in that category, based on fund performance and/or other fund characteristics (see, e.g., Kim, Shukla, and Tomas, 2000). This investor behavior results in a specific structure of mutual fund flows, that depends on a fund's relative performance within its category. As a consequence, the classification system also influences the incentives of fund managers, whose compensation is usually based on a percentage of fund assets (see, e.g., Khorana, 1996). Given that top performers in a category attract most of the inflows, fund managers have an incentive to maximize their performance relative to other funds in the same category. This may not be consistent with their shareholders' interests.

Currently, there are several coexisting categorization schemes of mutual funds. Traditionally, funds have been classified according to the investment *objective* declared in their prospectus, such as aggressive growth or growth-and-income. Along with the increase in the number of funds in 1990s, there appeared evidence that many funds exhibit investment behavior that cannot be characterized by their stated objective (see, e.g., Brown and Goetzmann, 1997, and diBartolomeo and Witkowski, 1997). In the early 1990s, several mutual fund data vendors such as Morningstar and Lipper introduced a new classification scheme, which was supposed to reflect funds' actual investment *style*. For example, Morningstar's equity style box is a three-by-three matrix with the division based on the market capitalization and book value of fund's latest portfolio. As witnessed in a recent Barra Strategic Consulting Group report (2001), "these style-based ranking systems have become well entrenched and highly influential" (p. 6). As one of the examples of this new trend, Business Week's Mutual Fund Scoreboard replaced in 1997 the stated objective classification system with a new one based on Morningstar's style box. Besides the stated objective and style categories, funds are often evaluated within broad investment *classes* such as "domestic stock" and "international stock". Such an approach is used to compute the widely publicized Morningstar star ratings. The class return rankings are often referred to in the financial press. For instance, the Wallstreet Journal Europe's Fund Scorecard regularly reports 15 leading and 10 lagging performers in the US equity, US bond, and other classes of funds.

Despite the wide use of these classification schemes, little is known about their relative impact on mutual fund flows. In the existing literature on the determinants of fund flows, two types of performance definitions are typically used: (i) in terms of raw or risk-adjusted returns (see, e.g., Gruber, 1996) and (ii) in terms of rankings within a stated objective category (see, e.g., Sirri and Tufano, 1998). All these studies find a clear positive relationship between fund flows and their past performance. Moreover, this relationship appears to be convex, i.e., the flow-performance sensitivity is higher for well-performing funds than for poorly performing funds (see, e.g., Chevalier and Ellison, 1997, and Sirri and Tufano, 1998). The lag structure of the flow-performance relationship is also nonlinear, with performance from six to eight months ago having the highest impact on current flows (see Chapter 3). Obviously, the use of raw or risk-adjusted returns ignores the effects of relative performance on fund flows. An alternative approach considers only the rankings based on the stated objective categorization scheme, assuming that investors only compare funds with the same investment objective. We would like to argue that ignoring the impact of alternative classification schemes on mutual fund flows, we are potentially missing important insights about investor behavior and the resulting incentives of fund managers.

In this chapter, we analyze the relationship between fund flows and their past relative performance with respect to different classification schemes. Specifically, we define a fund's relative performance as the (normalized) performance rank within three types of categories: funds with the same stated objective, funds with the same Morningstar style, and funds within the same asset class. Our primary goal is to disentangle the impact of fund relative performance on flows into the components corresponding to these *stated objective*, *Morningstar style*, and *asset class* categories. This will allow us to learn which mutual fund classification schemes are used by investors to form peer groups for the evaluation of fund performance. This information will be relevant for managers, who would like to know with which funds they should compete for investors' money. In addition, we investigate the relative impact of cardinal and ordinal performance measures on fund flows. This is important for determining whether fund managers are more interested in maximizing fund return or return ranking.

A related strand of literature analyzes the relationship between flows and performance on the level of fund family. It has been demonstrated that consumers invest more money in fund families with a star performer and fund families with higher average performance (see, e.g., Ivkovic, 2000). Moreover, Nanda, Wang, and Zheng (2000) find that top performance helps to boost flows not only to a star fund, but also to the other funds in the family. In this chapter, we study the star spillover effect in more detail. In particular, we investigate whether this effect differs across the family funds depending on their category (i.e., stated objective, Morningstar style, or asset class) and category of a star fund. The disentangling of the star spillover effect is important, since top performance of a star fund could actually "cannibalize" flows to the other funds in its family that have a similar investment approach, while boosting flows to the remaining funds in the family. In our analysis, we disentangle the flow spillover effect from a star fund to the other funds in the family into components corresponding to funds with the same stated objective, funds with the same Morningstar style, and funds in the same asset class as the star fund.

Our empirical results are based on the sample of US mutual funds belonging to one of four asset classes (domestic stock, international stock, taxable bond, and municipal bond) during the period from January 1993 (January 1994 for bond funds) to March 1999. In our main model, we analyze how monthly flows are influenced by the funds' normalized rankings based on the raw return over the last three years within the stated objective, Morningstar style, and asset class categories. We strongly reject the traditional specification, where relative performance is defined only with respect to the stated objective category. In all four fund classes, the Morningstar style and asset class rankings have both economically and statistically significant impact on fund flows on top of the impact of the stated objective rankings. With the only exception for international stock funds, the asset class ranking appears to be the most important relative performance measure for investors. It accounts for 58% of the total impact of relative performance on fund flows in the domestic stock class and about 80% and 86% in the taxable bond and municipal bond classes. Apparently, investors of these funds most often consider the asset class as a peer group for performance evaluation. Flows to domestic stock funds are also significantly related to their stated objective and Morningstar style rankings, which account for 28% and 14% of the total impact, respectively. The importance of the Morningstar style ranking is increasing over time, which reflects the growing interest to the Morningstar classification in the media as well as in the academic and applied literature (see, e.g., Bogle, 1998, and Davis, 2001). The stated objective and Morningstar style classification systems are especially important for institutional investors of domestic stock funds, who pay hardly any attention to the fund asset class rankings. In contrast, the differences between funds' objectives and styles do not play a large role for investors of taxable bond and municipal bond funds. The Morningstar style ranking accounts for about 10% of the total impact of relative performance on flows to bond funds, while the impact of the stated objective ranking is even smaller. Investors of international stock funds have a different hierarchy of the classification schemes, with 57% of the total impact of relative performance on flows due to the stated objective ranking and most of the remaining impact due to the asset class ranking. In contrast to the other asset classes, international stock funds are designed primarily for foreign investment and their region-based objectives are clearly defined. This may explain why the stated objective rankings of international stock funds appear reliable to their investors.

In a joint model of ordinal and cardinal performance measures, the impact of total return on fund flows never exceeds the combined impact of performance rankings. The ordinal measures of performance are especially important for investors of stock funds. In the domestic stock and international stock classes, the total impact of performance rankings on flows is approximately five times larger than the impact of total return. This implies that investors prefer to select funds that not only have a good performance, but also outperform their competitors.

The main model specification is extended to examine the star spillover effect in detail. We find that strong performance of a star fund (fund with the return over the last three years belonging to top 5% within its asset class) usually has a positive spillover effect on flows to the other funds in its family. For example, domestic stock nonstar funds from star families attract about 3.6% additional expected flows per year. In the municipal bond class, the flow spillover effect seems to be limited to funds with a similar investment approach (same Morningstar style or stated objective) as the star fund, which enjoy extra flows ranging from 2% to 5% per year. In the taxable bond class, the presence of a star fund is beneficial for flows to funds with the same stated objective as the star fund and detrimental for flows to funds with the same Morningstar style as the star fund, leading to the inflows of 6% and outflows of 4.5% per year, respectively. Naturally, star funds benefit the most from their stellar performance attracting as much as 18% extra flows in the domestic stock and municipal bond classes and about 9% extra flows in the taxable bond class. In the international stock class, we find no significant star spillover effect.

The structure of the chapter is as follows. Sections 4.2 and 4.3 describe the data set and the methodology, respectively. Section 4.4 presents the results concerning the relationship between fund flows and relative performance defined only with respect to the stated objective category. The results on the dependence of fund flows on performance rankings based on different classification schemes are discussed in Section 4.5. In Section 4.6, we investigate the relative impact of cardinal and ordinal measures of performance on fund flows. Section 4.7 is devoted to the category-specific analysis of the star spillover effect. In Section 4.8, we conclude and discuss the implications of our research.

4.2 Data description

In our empirical analysis, we use a merged data set taken from two sources: Micropal and Morningstar (April 1999 Principia Pro Data Disk). The former data set contains monthly total net asset values of US mutuals in January 1991 - March 1999, while the latter provides fund monthly returns in January 1970 - March 1999, inception date, annual equity and fixed-income style box classifications in 1992-1999, annual expense ratio and turnover rate in 1970-1999, and various fund characteristics (e.g., family indicator, front and deferred loads, 12b1 fees, manager tenure, minimum investment requirements, etc.) as of April 1999. The merged data set contains 9277 funds, which constitute approximately 87% funds covered by Morningstar in April 1999.

Currently, Morningstar classifies mutual funds into four investment classes: domestic stock, international stock, taxable bond, and municipal bond (money market funds are not included in the Morningstar's data base). As the flow-performance relationship is potentially different in the four classes, we estimate each model separately in the four samples corresponding to the fund classes. Since the domestic stock class is the largest and the most studied so far, it receives most of the attention in the discussion of the results. Since Morningstar started its style box classification in 1992, we take January 1993 - March 1999 as the sample period for domestic stock and international stock funds.[1] However, only a few bond funds have a fixed-income style data in 1992. Therefore, the sample period for taxable bond and municipal bond funds is from January 1994 to March 1999. Since we use a three-year period to evaluate fund performance, our analysis is restricted to the funds with at least three years of the return history. In order to reduce the impact of the typos and mergers, we exclude from our sample 1% of the outliers based on net relative flows (0.5% of the outliers with the largest positive flows and 0.5% of the outliers with the largest negative flows). The funds closed to the public and funds without Morningstar style data are also excluded from the sample.[2] In all regressions, flows are annualized to make our results comparable with the existing evidence.

It should be noted that our data set contains only the funds that survived till April

[1]Since Morningstar annual style classifications are as of the end of the year, we use fund style of year $t-1$ in the regressions of monthly flows realized in year t, assuming that the fund style did not change during the year. The results stay qualitatively the same, if fund style of year t is used in the regressions of that year.

[2]During the sample period, more than 94% of domestic and international stock funds had Morningstar equity style data, while about 67% of taxable and municipal bond funds had Morningstar fixed-income style data. The main results do not change, if we assign median style rankings to funds with missing style data and keep them in the sample.

Table 4.1: **Composition of the stated objective and Morningstar style categories of US domestic stock funds**

The table reports the number of fund-month observations with a given stated objective and Morningstar equity style accumulated within the domestic stock fund class over the period January 1993 - March 1999.

Objective / Style	Large value	Large blend	Large growth	Medium value	Medium blend	Medium growth	Small value	Small blend	Small growth
Aggressive growth	52	84	420	19	289	1959	84	268	1004
Growth	3200	7131	6854	2054	4680	5854	845	979	668
Growth-and-income	5858	9125	1234	1855	1161	169	35	88	24
Equity-income	3123	1427	80	939	283	36	45	0	0
Small company	0	0	0	218	516	1511	3444	3411	4590
Health	0	0	377	0	60	351	0	2	253
Financial	289	0	0	501	60	12	141	0	0
Natural resources	194	88	84	236	465	320	54	117	148
Precious metals	0	0	121	24	15	1182	51	56	376
Technology	36	48	297	48	72	795	0	24	103
Utility	1047	24	12	1764	15	0	3	0	0
Real estate	0	0	0	122	74	8	243	610	24
Communications	12	81	139	0	29	254	0	0	12
Asset allocation	901	2390	439	362	487	127	258	36	69
Balanced	3112	3465	1547	1340	780	493	175	16	58
Multi-asset global	492	699	86	325	327	238	116	121	48
Convertibles	451	147	120	136	290	90	75	107	20

1999 and is survivorship-biased. However, it is straightforward to show that this does not affect the consistency of OLS or WLS estimates, if past flows do not influence the probability of fund survival in a joint regression with returns. This assumption is in accordance with the empirical findings in Brown and Goetzmann (1995). Not surprisingly, Chevalier and Ellison (1997), Goetzmann and Peles (1997), and Sirri and Tufano (1998) find the same results in survivorship-biased and unbiased samples.

In order to illustrate the relation between the stated objective and Morningstar style classification schemes, Table 4.1 presents the composition of the respective categories in the domestic stock class. It shows the number of fund-month observations with a given stated objective and Morningstar equity style accumulated over the sample period. Domestic stock funds fall into one of five diversified stock objective categories (aggressive growth, growth, growth-and-income, equity-income, and small company), eight specialty stock objective categories (health, financial, natural resources, precious metals, technology, utility, real estate, and communications), and four hybrid objective categories (asset allocation, balanced, convertibles, and multi-asset global). These funds also belong to one of nine Morningstar equity style categories, which group funds on the basis of the market capitalization and growth potential of their portfolios (see Appendix 4.A for the definition of the Morningstar styles). Table 4.1 demonstrates that the objective and style classifications are not independent. For instance, about 75% of aggressive growth funds follow small and medium growth styles, while most of the funds with more conservative objectives (such as balanced and equity-income) are concentrated in the large value and large blend style categories. However, the dispersion in styles among the funds from the same objective category is quite high, which is consistent with the existing evidence on misclassification of funds in the objective categories (see, e.g., Brown and Goetzmann, 1997). Only in case of a few specialty stock objectives, there is a style containing more than 60% of the funds in a given objective category. In fact, eight from seventeen objective categories have funds spanning all nine cells in a style box. Similar levels of dispersion across styles are also observed in the other fund classes. As a result, there is sufficient variation between fund rankings relative to the stated objective and Morningstar style categories, which allows me to identify their separate effects on fund flows.

Table 4.2 presents the summary statistics of the funds belonging to the four classes under consideration, calculated throughout the sample period. During the sample period, an average domestic stock fund had a thirteen-year performance record, $766 million in total net assets and 1.3% expense ratio and experienced approximately 8% net flow per year. In contrast, funds from the other asset classes were smaller (from $289 million to $544 million in total net assets for an average municipal bond and international stock fund, respectively) and younger (7-9 years) and attracted lower net flows (about 1% per year) during the sample period. Since most families include different types of funds, the family characteristics are similar across the classes. An average fund family includes about 14 funds and controls approximately $4.5 billion of assets.

Table 4.2: **Summary statistics of US mutual funds**

The table presents summary statistics of the domestic stock, international stock, taxable bond, and municipal bond classes of US mutual funds. For each fund class, the table reports the mean and standard deviation of the respective fund characteristics over the sample period (January 1993 - March 1999 for stock funds and January 1994 - March 1999 for bond funds). Note that total return and flows are annualized.

Fund characteristics	Dom. stock		Int. stock		Tax. bond		Mun. bond	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Net relative flow, %	7.96	55.59	1.07	69.4	1.62	60.51	0.91	38.57
Total net assets, $mln	766.37	2294.32	544.43	1648.07	431.06	987.59	288.82	798.49
Age, years	12.52	13.65	6.94	5.92	8.83	8.06	7.39	4.06
Expense ratio, %	1.32	0.57	1.73	0.62	1.02	0.48	0.94	0.42
Total risk, %	3.51	1.18	4.17	1.29	1.19	0.62	1.34	0.3
# funds in the family	13.35	26.58	14.08	28.3	13.83	27.74	13.91	27.92
Family TNA, $bln	4.47	20.30	4.76	22.11	4.58	20.99	4.63	21.28
Family age, years	15.11	17.68	15.07	17.58	15.06	17.62	15.06	17.6
Objective flow, %	6.04	15.11	3.14	14.28	-0.98	13.76	-1.97	3.36
Style flow, %	6.23	10.01	1.81	15.63	-0.22	11.87	-2.58	6.45
Total return, %	16.47	5.49	8.45	7.79	7.22	2.3	6.45	0.93

4.3 Methodology

In this chapter, we analyze the relationship between fund flows and their past relative performance with respect to three types of categories: *stated objective*, *Morningstar style*, and *asset class*. For a given fund, these categories consist of funds with the same stated objective, funds with the same Morningstar style (equity style for stock funds and fixed-income style for bond funds), and funds within the same Morningstar asset class (domestic stock, international stock, taxable bond, or municipal bond), respectively. Note that the stated objective and Morningstar style categories include only the funds within the respective asset class, since we do not consider performance relative to funds from the other asset classes. Estimating the flow-performance relationship, we would like to control for the other determinants of flows previously identified in the literature, such as size, age, fees, and risk of a given fund, size, age and number of funds in the fund's family, and category-specific flows (see, e.g., Sirri and Tufano, 1998, and Nanda, Wang, and Zheng, 2000). Since some of these factors were also found to affect the sensitivity of flows with respect to performance (see, e.g. Chevalier and Ellison, 1997, and Sirri and Tufano, 1998), we model both the performance-unrelated part of flows and flow-performance sensitivity as linear functions of the control variables. Specifically, our basic model is as follows:

$$f_{i,t} = x'_{i,t-1}a + (x'_{i,t-1}b) * (c_1 RP^{obj}_{i,t-1} + c_2 RP^{style}_{i,t-1} + c_3 RP^{class}_{i,t-1}) + e_{i,t}, \tag{4.1}$$

where $RP^{obj}_{i,t-1}$, $RP^{style}_{i,t-1}$, and $RP^{class}_{i,t-1}$ denote fund i's relative performance with respect to the stated objective, Morningstar style, and asset class categories and $x'_{i,t-1}$ is a vector of fund i's control variables at month $t-1$. In order to identify the parameters in the model, we impose the restriction that the sum of the performance coefficients c_k be equal to one.

The dependent variable in the model is fund i's net relative flow. In line with the previous studies (see, e.g., Gruber, 1996), it is defined as the growth in the fund assets net of reinvested dividends:

$$f_{i,t} = \frac{TNA_{i,t} - (1 + R_{i,t})TNA_{i,t-1}}{TNA_{i,t-1}} \tag{4.2}$$

where $TNA_{i,t}$ denotes fund i's total net assets at the end of month t and $R_{i,t}$ is return of fund i over month t. Here, we assume that all earnings are automatically reinvested in the fund and that flows occur at the end of the month.

Fund i's performance relative to a given category RP^{cat}_i is measured as a *normalized category ranking*:

$$RP^{cat}_{i,t} = [\rho^{cat}_{i,t} - \frac{1}{2}] \frac{p(.95)^{class}_{i,t} - p(.05)^{class}_{i,t}}{p(.95)^{cat}_{i,t} - p(.05)^{cat}_{i,t}}, \tag{4.3}$$

where $\rho^{cat}_{i,t}$ is fund i's fractional rank in a given category based on its total return over the past 36 months[3] and $p(x)^{cat}$ is x^{th} percentile of three-year returns of funds belonging to a given category. Defining fund ranking as a fractional rank (a fraction of funds from the category with lower return than fund i) ranging from 0 to 1 is traditional in the literature (see e.g. Sirri and Tufano, 1998). We adjust the fractional rank by 0.5

[3]The three-year performance horizon should be long enough, since in Chapter 3 we found that past three years account for approximately 90% of the total performance impact on flows. We use total returns to construct rankings, since fund rankings published in the mass media are typically based on raw rather than risk-adjusted returns. We account for the effect of fund risk on flows by including a measure of a fund's total risk as one of the control variables.

to assign zero ranking to a median fund in the category. This allows us to interpret the performance-unrelated part $x'_{i,t-1}a$ of the model as expected flows of a fund with neutral performance. The original category rankings are normalized to make sure that they change by the same amount in response to a given change in fund's return.[4] For a given category, the normalization coefficient is equal to the return dispersion in the asset class divided by the return dispersion in this category.[5] Thus, the asset class rankings do not change with normalization, while the normalized stated objective and Morningstar style rankings become equivalent to the asset class rankings. After normalization, the total impact of a unit change in fund's relative performance on its flows is equal to the flow-performance sensitivity times the sum of the performance coefficients. Because of the identifying restriction that the sum of the performance coefficients be equal to one, each performance coefficient c_k can be interpreted as a percentage of the total impact of relative performance on fund's flow due to the change in rankings in the respective category.

Both the performance-unrelated flows and flow-performance sensitivity are modelled as linear functions of the control variables: $x'_{i,t-1}a$ and $x'_{i,t-1}b$, respectively. By construction, a parameters measure the impact of the control variables on expected flows of a fund with median category rankings, while b parameters show how the sensitivity of flows to performance varies with fund characteristics. The vector of fund i's control variables $x'_{i,t-1}$ includes a constant, four fund-specific factors (log size, log age, expense ratio, and total risk of a fund), three family-specific factors (log of the number of funds in the family, log size and log age of the family), and two category-specific factors (stated objective and Morningstar style category flows).

The effects of fund size and age on flows were identified, e.g., in Chevalier and Ellison (1997) as well as in Chapter 3, where it is demonstrated that smaller and younger funds

[4] A simple example illustrates that a direct comparison of the original category ranking coefficients can be misleading. Suppose that 1% change in return is equivalent to 0.1 change in the objective category rankings and 0.3 change in the style category rankings and that the estimated ranking coefficients are 10 and 5, respectively. However, the fact that the first coefficient is larger than the second does not imply that flows are more sensitive to the objective category rankings. A 1% change in fund return would lead to 10∗0.1=1% change in flows due to the objective category rankings and 5∗0.3=1.5% change in flows due to the style category rankings.

[5] A return dispersion in a given category is defined as the difference between 95% and 5% return percentiles in the category rather than the difference between the maximum and minimum returns in the category, since it is more robust against the outliers.

enjoy larger performance-unrelated flows as well as higher flow-performance sensitivity. These findings could be caused by more active advertising by young funds, who have not yet acquired the long-term reputation of the old funds. We include log fund size and log fund age to the model to control for these effects. The impact of the expense ratio on flows can be twofold. On the one hand, higher expense ratio may lead to lower flows, as investors would like to maximize net-of-fee earnings. On the other hand, a higher expense ratio is associated with larger advertising expenditures and may increase fund flows. The existing evidence does not give a clear answer which effect prevails. Most studies document a negative impact of fund expense ratios on flows (see, e.g., Nanda, Wang, and Zheng, 2000), while Barber, Odean, and Zheng (2001) find a positive relation between expenses and flows in a sample of large diversified equity funds. If investors prefer less risk, flows should be negatively related to fund total risk measured as the standard deviation of total returns during the last 36 months. Indeed, several studies find a negative relation between fund flows and total risk (see, e.g., Nanda, Wang, and Zheng, 2000, and Barber, Odean, and Zheng, 2001).

The dependence of individual fund flows on family characteristics has received much attention in the recent literature (see, e.g., Nanda, Wang, and Zheng, 2000). It has been documented that funds belonging to larger families attract higher flows (see Ivkovic, 2000), while older families achieve larger market share (see Khorana and Servaes, 2001). These findings are probably due to the greater visibility and better distribution networks available to the funds from large and old families. We measure family size by log of the number of funds in the family and log of the family total net assets. Family age is measured as log of the age of the oldest fund in the family.

The existing studies document a strong dependence of individual fund flows on flows to fund's objective category (see, e.g., Sirri and Tufano, 1998). In our model, we include both stated objective and Morningstar style category flows as the control variables. They should control for the temporal changes in the individual fund flows due to movement in the category-specific flows. The objective and style flows are measured as the TNA-weighted average net relative flows of funds within the respective categories.

Throughout the chapter, we run panel regressions with month dummies whose coefficients are assumed to lie on a polynomial of order p:

$$Dummy_t = \sum_{k=0}^{p} \theta_k t^{-k}, \qquad t = 1, ..., T. \tag{4.4}$$

The empirical results indicate that $p = 5$ suffices (the higher-order θ's are insignificant and their inclusion does not influence our findings). We compute weighted least squares estimates with the variance of the residuals $e_{i,t}$ modelled as

$$Var(e_{i,t}) = \exp(x'_{i,t-1}w), \tag{4.5}$$

where $x'_{i,t-1}$is the vector of fund i's control variables defined as before. The w coefficients are estimated on the basis of the OLS residuals. The model parameters in (4.1) are estimated by means of a concentrated least-squares approach. For the pre-specified values of b parameters, the model (4.1) is linear in the remaining parameters. Therefore, one can easily calculate the least squares estimates of the remaining parameters and the corresponding sum of the squared residuals. By numerically maximizing the concentrated sum of squares over the b parameters, we obtain the least squares estimates of the b parameters and, consequently, of the remaining parameters in the model.

4.4 Impact of performance relative to the stated objective category

In the previous studies of the flow-performance relationship, fund's relative performance is typically measured with respect to its stated objective category (see, e.g., Sirri and Tufano, 1998). Therefore, we use the normalized stated objective rankings RP^{obj} as the only performance measure in our first model specification, which is equivalent to model (4.1) with the additional restriction $c_2 = c_3 = 0$.

Table 4.3 reports the results based on samples of domestic stock, international stock, taxable bond, and municipal bond funds. During the sample period, an average domestic stock, international stock, or municipal bond fund with median performance relative to its objective category attracted between 5% and 7% net flows per year. Flows to taxable bond funds were much lower, about 1.3% per year. Consistent with the existing evidence, we find a clear positive relation between the stated objective ranking of a fund and its flows. A 10 percentile move in the stated objective ranking leads to about 3% additional flows for an average domestic or international stock fund. Flows to taxable bond and municipal bond funds would change by approximately 1.3% and 1.6%, respectively.

Most of the control variables prove to have both economically and statistically significant impact on expected fund flows. Confirming the presence of the size effect, we find

Table 4.3: **Relationship between flows and relative performance with respect to the stated objective category**

The table documents the relationship between fund flows and stated objective rankings in the domestic stock, international stock, taxable bond, and municipal bond classes of US mutual funds. For each fund class, the table reports the estimated coefficients and standard errors (in the parentheses) based on the model (4.1) with the restriction $c_2 = c_3 = 0$. The identifying restriction is that the coefficient on the stated objective ranking c_1 is equal to one. The dependent variable is fund's annualized net relative flow. The control variables include a constant, logs of fund size and age, fund expense ratio and risk, log of number of funds in the family, logs of fund family size and age, objective and style category flows. Both the performance-unrelated flows and flow-performance sensitivity are linear functions of the control variables. The sample period is January 1993 - March 1999 for stock funds and January 1994 - March 1999 for bond funds.

	Dom. stock		Int. stock		Tax. bond		Mun. bond	
Performance-unrelated flows								
$Const$	11.32	(2.91)	25.51	(6.22)	3.28	(4.17)	10.45	(3.26)
$\log FundTNA_{i,t-1}$	-2.42	(0.15)	-2.62	(0.55)	-2.15	(0.35)	-1.25	(0.19)
$\log FundAge_{i,t-1}$	-4.63	(0.31)	-6.37	(0.97)	-4.59	(0.84)	-6.50	(0.79)
$FundExpenseRatio_{i,t-1}$	3.21	(0.61)	-3.66	(1.04)	-0.79	(0.84)	-0.81	(1.04)
$FundRisk_{i,t-1}$	-1.51	(0.29)	-0.35	(0.50)	-0.19	(0.90)	-2.24	(0.87)
$\log Family\#funds_{i,t-1}$	-3.46	(0.28)	-3.33	(0.75)	-4.34	(0.46)	-5.14	(0.32)
$\log FamilyTNA_{i,t-1}$	2.88	(0.24)	2.96	(0.64)	3.97	(0.54)	3.38	(0.24)
$\log FamilyAge_{i,t-1}$	2.75	(0.35)	-1.01	(0.67)	-0.92	(0.75)	-0.73	(0.31)
$ObjectiveFlow_{i,t-1}$	0.42	(0.10)	0.88	(0.03)	0.53	(0.09)	0.35	(0.05)
$StyleFlow_{i,t-1}$	0.36	(0.04)	0.29	(0.04)	0.16	(0.05)	0.42	(0.06)
Flow-performance sensitivity								
$Const$	-5.77	(3.61)	65.13	(9.30)	1.22	(3.96)	34.14	(6.72)
$\log FundTNA_{i,t-1}$	-0.16	(0.42)	1.94	(0.76)	0.50	(0.38)	0.13	(0.53)
$\log FundAge_{i,t-1}$	-4.35	(0.58)	-3.17	(1.64)	-3.93	(1.33)	-5.13	(2.09)
$FundExpenseRatio_{i,t-1}$	7.65	(1.09)	-0.80	(1.66)	7.80	(1.28)	-2.69	(3.42)
$FundRisk_{i,t-1}$	6.54	(0.94)	-5.81	(1.02)	1.81	(1.00)	-7.62	(1.96)
$\log Family\#funds_{i,t-1}$	-4.79	(0.67)	8.18	(1.17)	-2.74	(0.60)	6.72	(1.08)
$\log FamilyTNA_{i,t-1}$	1.87	(0.66)	-4.67	(1.00)	3.06	(0.72)	-1.77	(0.63)
$\log FamilyAge_{i,t-1}$	3.98	(0.67)	-2.54	(1.40)	-3.25	(1.17)	-2.46	(1.17)
$ObjectiveFlow_{i,t-1}$	0.15	(0.14)	0.13	(0.08)	0.13	(0.06)	0.00	(0.11)
$StyleFlow_{i,t-1}$	0.04	(0.06)	-0.20	(0.05)	-0.06	(0.03)	-0.04	(0.14)

that a twofold increase in fund size is associated with a flow decrease from about 0.9% for a municipal bond fund to about 1.8% for a domestic or international stock fund. The sensitivity of fund flows to performance is only marginally (and in most classes insignificantly) affected by fund size. As expected, younger funds enjoy larger flows as well as higher flow-performance sensitivity. All other things being equal, a two-time difference in age implies approximately 4% difference in flows of a fund with median performance and 0.3% difference in the sensitivity of fund flows with respect to 10 percentile change in the objective rankings.

The effect of the expense ratio on flows differs across the asset classes. In response to a 1% increase in the expense ratio, flows to a domestic stock fund will increase by about 3.2%, while flows to an international stock fund will decrease by approximately 3.7%. In case of bond funds, the expense ratio also has a negative although insignificant impact on flows. Our finding of a positive relation between expenses and flows of domestic stock funds is similar to that of Barber, Odean, and Zheng (2001), who explain it by greater marketing efforts of funds with higher expense ratios. It appears that this effect does not outweigh the costs associated with higher fees for investors of the other fund classes. Probably, larger advertising expenditures of domestic stock and taxable bond funds with higher expense ratios make their flows more sensitive to past performance. In these classes, a 1% increase in the expense ratio is associated with approximately 0.7% rise in sensitivity of flows to 10 percentile move in the stated objective rankings.

In line with the expectations, we find a negative relation between fund flows and total risk, which is significant in the domestic stock and municipal bond classes. In these classes, a 1% increase in fund's total risk leads on average to about 1.5% and 2.2% decrease in the subsequent flows. The effect of total risk on the flow-performance sensitivity is more ambiguous, being significantly positive for domestic stock funds and significantly negative for international stock and municipal bond funds.

Consistent with the previous studies (see, e.g., Nanda, Wang, and Zheng, 2000), we find a significant impact of the family-specific variables on individual fund flows. In all four classes, funds belonging to families with larger total net assets attract significantly higher flows. At the same time, the growth of fund families may have a cost, since the number of funds in a family appears to have a significantly negative impact on flows. This is probably explained by investors' preference to invest in more focused families. On average, a twofold increase in family size or a twofold decrease in the number of funds in the family yields about 2-3% additional flows. The impact of the age of the fund family on flows is significantly positive in the domestic stock class and significantly negative in the municipal bond class (in the other classes, it is also negative, but insignificant). A two-time difference in fund family age implies about 2% and 0.5% difference in flows to funds belonging to these classes, respectively. The effects of the family-specific variables on the flow-performance sensitivity differ across the fund classes. Flows to an average domestic stock fund are more sensitive to its past performance, if it belongs to a larger or older family or family with the lower number of funds.

Finally, we find that fund flows are strongly related to the category-specific flows. A domestic stock fund with a median stated objective ranking is expected to attract about 42% of its objective flow and 36% of its style flow. Flows to international stock and taxable bond funds seem to be even more related to the objective flows, while flows to municipal bond funds are more sensitive to the style flows. These results provide preliminary evidence of the impact of the Morningstar style classification scheme on fund flows, which appear to be related not only to the objective-specific, but also to the style-specific flows. Further analysis of the category-specific factors driving fund flows is carried out in the next section.

4.5 Impact of performance relative to the stated objective, Morningstar style, and asset class categories

The main goal of the present chapter is to examine the dependence of fund flows on relative performance measures based on the alternative classification schemes. Therefore, in this section we relax the assumption that fund flows are driven only by the stated objective rankings and analyze the relationship between fund flows and their normalized performance rankings within the stated objective, Morningstar style, and asset class categories.

Table 4.4 presents the results based on model (4.1). The coefficients on the control variables are hardly affected by the introduction of the additional performance measures and are omitted from the table. Panel A of Table 4.4 reports the performance coefficients and the corresponding standard errors, while Panel B of Table 4.4 presents p-values of

Table 4.4: **Relationship between flows and relative performance with respect to the stated objective, Morningstar style, and asset class categories**
The table documents the relationship between fund flows and stated objective, Morningstar style, and asset class rankings (RP^{obj}, RP^{style}, and RP^{class}, respectively) in the domestic stock, international stock, taxable bond, and municipal bond classes of US mutual funds. For each fund class, the table presents the results based on the model (4.1). The identifying restriction is that the sum of the performance coefficients is equal to one. The dependent variable is fund's annualized net relative flow. The control variables include a constant, logs of fund size and age, fund expense ratio and risk, log of number of funds in the family, logs of fund family size and age, objective and style category flows. Both the performance-unrelated flows and flow-performance sensitivity are linear functions of the control variables. The sample period is January 1993 - March 1999 for stock funds and January 1994 - March 1999 for bond funds. Since the coefficients on the control variables stay approximately the same as in the previous model (see Table 4.3), Panel A of the table only reports the estimated coefficients and standard errors (in the parentheses) of the performance variables. Panel B presents p-values of the pairwise tests for the equality of the respective performance coefficients.

Panel A. Performance coefficients

	Dom. stock		Int. stock		Tax. bond		Mun. bond	
Relative performance measures								
RP^{obj}	0.28	(0.04)	0.57	(0.06)	0.09	(0.04)	0.04	(0.18)
RP^{style}	0.14	(0.03)	0.09	(0.07)	0.12	(0.03)	0.10	(0.05)
RP^{class}	0.58	(0.05)	0.34	(0.11)	0.80	(0.05)	0.86	(0.21)

Panel B. P-values of the pairwise tests

	Dom. stock	Int. stock	Tax. bond	Mun. bond
H_0: $c_1 = c_2$ (RP^{obj} vs. RP^{style})	0.013	0.000	0.456	0.725
H_0: $c_1 = c_3$ (RP^{obj} vs. RP^{class})	0.000	0.171	0.000	0.038
H_0: $c_2 = c_3$ (RP^{style} vs. RP^{class})	0.000	0.147	0.000	0.002

the pairwise tests of equality between the performance coefficients.

First of all, we find that the Morningstar style and asset class rankings have a strong positive impact on fund flows on top of the impact of the stated objective rankings. In all four asset classes, we strongly reject the traditional specification, where relative performance is defined only with respect to the objective category. In fact, the asset class ranking appears to be the most important relative performance measure for investors of domestic stock as well as investors of taxable and municipal bond funds. In these asset classes, the difference between the coefficients on the asset class ranking and the other two types of rankings is significant at 1% level (see Table 4.4, Panel B). Only in the international stock class, fund flows are most sensitive to the stated objective ranking, although the difference between its coefficient and coefficient on the asset class ranking is not significant.

By construction, each performance coefficient can be interpreted as the percentage of the total impact of relative performance on fund flows attributed to the respective category. Alternatively, one can also interpret the performance coefficients as the percentage of investors, who use categories based on the respective classification schemes as peer groups for the evaluation of fund performance. The relative impact of the stated objective, Morningstar style, and asset class rankings on fund flows in different fund classes is illustrated graphically in Figure 4.1. In the domestic stock class, 58% of the total impact is attributed to the asset class ranking, while about 28% and 14% of the total impact are due to the stated objective and Morningstar style rankings, respectively. Thus, there is a clear hierarchy of the different classification schemes with respect to their importance for investors. Apparently, investors of domestic stock funds (or their financial advisors) most often consider the asset class as a peer group for performance evaluation, followed by the stated objective category and Morningstar style category. This is consistent with the anecdotal evidence from one of managing directors of The Vanguard Group that their equity fund investors most often ask about the best-performing equity funds rather than, say, the best-performing growth funds. This is also consistent with Carhart, et al. (2002), who find that marking up at the last trading day of the year to improve a calendar-year performance is stronger among equity funds with top "universe-relative performance" than among equity funds with top "category-relative performance" (in our terminology, funds with top asset class rankings and funds with top stated objective rankings, respectively).

Investors of taxable bond and municipal bond funds use a similar hierarchy, in which the asset class ranking is even more important, accounting for 80% and 86% of the total impact of relative performance on flows, respectively. The differences in fund investment policies do not seem to play a large role for these investors, which may explain the marginal impact of the stated objective and Morningstar style rankings on flows to bond funds. Investors of international stock funds appear to have a different hierarchy of the classification schemes, with 57% of the total impact of relative performance on flows due to the stated objective ranking and 34% due to the asset class ranking (the Morningstar style ranking has a positive but insignificant impact on flows). In contrast to the other fund classes, international stock funds are designed primarily for foreign investment. Typically, their objective is to invest in stocks from a certain geographical

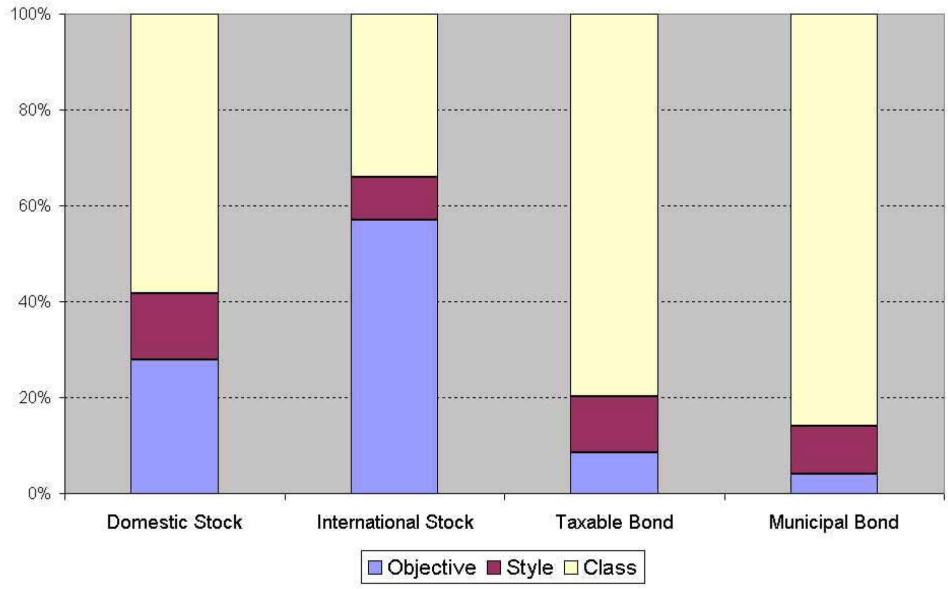

Figure 4.1: **Relative impact of different classification schemes on flows to US mutual funds**

The graph presents the coefficients of the stated objective, Morningstar style, and asset class rankings for the domestic stock, international stock, taxable bond, and municipal bond classes of US mutual funds, based on model (4.1). Each performance coefficient is interpreted as a percentage of the total impact of fund relative performance on flows, which is due to the respective classification scheme. The sample period is January 1993 - March 1999 for stock funds and January 1994 - March 1999 for bond funds.

region such as Latin America or Japan. Since these region-based objectives are more clearly defined and also appear more distinct from each other than, say, growth-and-income and equity-income objectives in the domestic stock class, the stated objective rankings of international stock funds may seem reliable to their investors.

The hierarchy of the different classification schemes for investors of domestic stock funds is further illustrated in Figure 4.2. It presents the relative impact of the stated objective, Morningstar style, and asset class rankings on fund flows within different subsamples of the domestic stock class. First, we discuss the results based on the periods January 1993 - December 1996 and January 1997 - March 1999. We observe that the stated objective as well as Morningstar style rankings have become somewhat more important over time. Probably, the latter tendency is due to the fact that in 1997 Morningstar started to use its equity and fixed-income style boxes to categorize funds. Note that Morningstar category does not always coincide with Morningstar style. The former is based on fund's investment policy over the past three years, while the latter is based on fund's latest portfolio holdings. In addition, Morningstar keeps using the stated objective categories for specialty funds and classifies most hybrid funds into the domestic hybrid or international hybrid categories. The highest correlation between Morningstar styles and categories is for the diversified stock funds, which constitute the majority of the domestic stock class. Indeed, the concept of Morningstar equity styles is most applicable to this type of funds, since hybrid funds invest a large part of their portfolios in bonds and specialty stock funds invest in narrow market segments. This may explain why the impact of the Morningstar style ranking is higher for diversified stock funds than for specialty stock and hybrid funds. In general, the hierarchy of the classification schemes for diversified stock, specialty stock, and hybrid funds is similar to that of all domestic stock funds, with the asset class ranking having the largest impact on flows. The asset class ranking is especially important for investors of specialty stock funds, who seem to attach only marginal importance to funds' stated objective rankings that are based on a relatively small number of funds specializing in the same market sector.

There are remarkable differences between the hierarchies of the classification schemes in the subsamples of primarily private and primarily institutional funds (defined in line with Chevalier and Ellison, 1997, as funds with a minimum initial purchase of less than $25,000 and at least $25,000, respectively). Most private investors compare

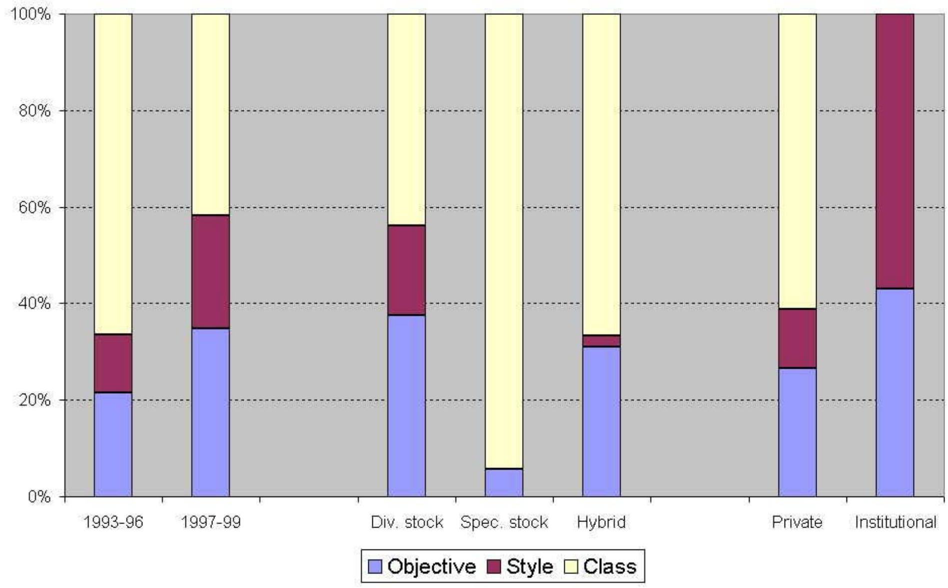

Figure 4.2: **Relative impact of different classification schemes on flows to domestic stock funds**

The graph presents the coefficients of the stated objective, Morningstar style, and asset class rankings on flows to the different subsamples of the domestic stock class, based on model (4.1). Each performance coefficient is interpreted as a percentage of the total impact of fund relative performance on flows, which is due to the respective classification scheme. Unless otherwise specified, the sample period is January 1993 - March 1999. The first two columns in the graph present results based on the periods January 1993 - December 1996 and January 1997 - March 1999. The second division of the domestic stock class into subsamples (see columns three to five) is based on fund stated objectives. Diversified stock funds have an aggressive growth, growth, growth and income, equity-income, or small company objective, specialty stock funds have a health, financial, natural resources, precious metals, technology, utility, real estate, or communications objective, while hybrid funds have an asset allocation, balanced, convertibles, or multi-asset global objective. The last two columns in the graph are based on the subsamples of primarily private and primarily institutional funds, defined as funds with a minimum initial purchase of less than $25,000 and at least $25,000, respectively.

return of a given fund with that of the other domestic stock funds. This indicates that they are style timers, choosing funds with larger holdings of recent winners. In contrast, institutional investors use only the categories grouping funds with a similar investment approach for performance evaluation, attaching approximately equal weights to the stated objective and Morningstar style classifications. This finding is similar to that of Del Guercio and Tkac (2002) who document that pension fund clients (primarily institutional investors) use more quantitatively sophisticated performance measures than mutual fund customers (primarily private investors).

A number of existing studies (see, e.g., Sirri and Tufano, 1998) have demonstrated the nonlinearity of the flow-performance relationship. To adjust for this, we extended the basic model specification (4.1) to allow the coefficients on the stated objective, Morningstar style, and asset class rankings differ across the respective performance quintiles. In all four asset classes, the average impact of the different relative performance measures based on the extended model remained practically the same as in the basic model (4.1). These results are not reported here, but are available upon request.

4.6 Impact of ordinal as well as cardinal measures of performance

So far, the analysis has been limited to the ordinal measures of performance based on the three classification schemes under consideration. All these measures are based on fund's total return, which is a cardinal performance measure. It is an interesting question which type of performance measure, cardinal or ordinal, is more important for investors. To formulate it differently, do investors pay more attention to fund returns or return ranks? To answer this question, we extend the basic model (4.1) by adding a cardinal performance measure to the regressors:

$$f_{i,t} = x'_{i,t-1}a + (x'_{i,t-1}b) * (c_1 RP^{obj}_{i,t-1} + c_2 RP^{style}_{i,t-1} + c_3 RP^{class}_{i,t-1} + c_4 return_{i,t-1}) + e_{i,t}, \quad (4.6)$$

where $return_{i,t-1}$ is fund i's total return over the last 36 months (from $t-1$ to $t-36$). As before, the identifying restriction is that the sum of c_1, c_2, and c_3 be equal to one.

Table 4.5 reports the estimation results (as before, we omit the coefficients on the control variables, since they stay approximately the same). In a joint model, both ordinal

Table 4.5: **Relationship between flows and ordinal as well as cardinal performance measures**

The table documents the relationship between fund flows and ordinal as well as cardinal performance measures in the domestic stock, international stock, taxable bond, and municipal bond classes of US mutual funds. The ordinal performance measures are fund's stated objective, Morningstar style, and asset class rankings (RP^{obj}, RP^{style}, and RP^{class}, respectively), while the cardinal performance measure is fund's total return over the last 36 months. For each fund class, the table presents the results based on the model (4.6). The identifying restriction is that the sum of the coefficients on the ordinal performance measures is equal to one. The dependent variable is fund's annualized net relative flow. The control variables include a constant, logs of fund size and age, fund expense ratio and risk, log of number of funds in the family, logs of fund family size and age, objective and style category flows. Both the performance-unrelated flows and flow-performance sensitivity are linear functions of the control variables. The sample period is January 1993 - March 1999 for stock funds and January 1994 - March 1999 for bond funds. Since the coefficients on the control variables stay approximately the same as in the previous model (see Table 4.3), the table only reports the estimated coefficients and standard errors (in the parentheses) of the performance variables. The last row of the table presents the normalized coefficient on total return, which is equal to the original coefficient on total return times the average return dispersion (the difference between the maximum and minimum returns) in a given class.

	Dom. stock		Int. stock		Tax. bond		Mun. bond	
Ordinal performance measures								
RP^{obj}	0.31	(0.05)	0.64	(0.08)	0.15	(0.07)	0.16	(0.25)
RP^{style}	0.15	(0.03)	0.08	(0.08)	0.19	(0.05)	0.15	(0.05)
RP^{class}	0.54	(0.06)	0.27	(0.12)	0.66	(0.10)	0.69	(0.27)
Cardinal performance measure								
$return$	0.010	(0.002)	0.008	(0.007)	0.097	(0.028)	0.146	(0.054)
$return * dispersion$	0.182		0.189		0.951		0.507	

and cardinal performance measures prove to be significant. In all asset classes except for the international stock class, fund's return has a significantly positive impact on flows. For instance, in the domestic stock class a 1% change in return keeping all performance rankings unchanged is associated with 0.4% change in fund flows. The inclusion of the cardinal performance measure hardly changes the category-specific composition of the total impact of ordinal performance measures on flows. In the domestic stock, taxable bond, and municipal bond classes, the asset class ranking remains the most important relative performance measure, although its weight has slightly decreased. In order to compare the relative impact of ordinal and cardinal performance measures on fund flows, we normalize the return coefficient by multiplying it by the average return dispersion (the difference between the maximum and minimum returns) in the respective asset class (see the last row of Table 4.5). For the domestic stock and international stock funds, the normalized return coefficient is about 0.18, which implies that the impact of total return on fund flows is about 18% of the combined impact of relative performance measures. The total return is relatively more important for investors of municipal bond and especially taxable bond funds, accounting for about 51% and 95% of the total impact of relative performance, respectively. These results suggest that ordinal performance measures are at least as important for investors as cardinal performance measures. In other words, investors (especially those of stock funds) prefer to select funds that not only have a good performance but also outperform the others.

4.7 Category-specific flow spillover effect

In this section, we examine whether the star spillover effect identified, e.g., in Nanda, Wang, and Zheng (2000) differs across the funds from star families depending on their category (i.e., stated objective, Morningstar style or asset class) and category of a star fund. Specifically, we disentangle the flow spillover effect from a star fund to the other funds in the family into components corresponding to funds with the same stated objective, funds with the same Morningstar style, and funds within the same asset class as the star fund. In line with the previous studies (see, e.g., Ivkovic, 2000), we define a star fund as a fund with the performance (the total return over the past 36 months) in top 5% within its asset class. The basic model (4.1) is extended as follows:

$$f_{i,t} = x'_{i,t-1}a + (x'_{i,t-1}b) * (c_1 RP^{obj}_{i,t-1} + c_2 RP^{style}_{i,t-1} + c_3 RP^{class}_{i,t-1} \\ + d_1 D^{star}_{i,t} + d_2 D^{obj}_{i,t} + d_3 D^{style}_{i,t} + d_4 D^{class}_{i,t}) + e_{i,t}, \tag{4.7}$$

where $D^{star}_{i,t}$ is a dummy equal to one if fund i is a star fund and $D^{obj}_{i,t}$ ($D^{style}_{i,t}$, $D^{class}_{i,t}$) is a dummy equal to one if fund i belongs to the same stated objective (Morningstar style, asset class) category as one of the star funds in its family. Since within each asset class $D^{class}_{i,t}$ is equal to one for all funds in a star family (family with at least one star fund), d_4 shows the difference between expected flows to funds from star families and other funds. For funds from star families, the coefficients d_1, d_2, and d_3 show the additional expected flows of a star fund, funds with the same stated objective as one of the star funds, and funds with the same Morningstar style as one of the star funds, respectively. As before, we impose the identifying restriction that the sum of c_1, c_2, and c_3 be equal to one.

The estimation results are presented in Table 4.6 (as before, the coefficients on the control variables are omitted). Consistent with Ivkovic (2000), we find a presence of the statistically and economically significant flow spillover effect from a star fund to the other funds in the family in all asset classes except for the international stock class. The flow spillover effect is typically positive, as the presence of a star fund helps to boost flows to the other funds in the family. In the domestic stock class, the magnitude of the flow spillover effect does not differ significantly (at the 5% level) across different types of nonstar funds in star families that attract about 3.6% additional expected flows per year. In the municipal bond class, the flow spillover effect seems to be limited to funds with the same Morningstar style and funds with the same stated objective as one of the star funds in their families, which enjoy about 5% and 2% extra flows, respectively. In the taxable bond class, the presence of a star fund is beneficial for flows to funds with the same stated objective as the star fund and detrimental for flows to funds with the same Morningstar style as the star fund, leading to the inflows of 6% and outflows of 4.5% per year, respectively. Naturally, star funds benefit the most from their stellar performance attracting as much as 18% extra flows in the domestic stock and municipal bond classes and about 9% extra flows in the taxable bond class.

Table 4.6: **Category-specific flow spillover effect from a star fund to the other funds in the family**

The table documents the flow spillover effect from a star fund to the other funds in the family, including funds with the same stated objective, funds with the same Morningstar style, and funds within the same asset class (see the coefficients on D^{obj}, D^{style}, and D^{class}, respectively). The coefficient on D^{star} captures the difference between flows to a star fund and flows to the other funds in the star family. The corresponding results are based on the model (4.7) and samples of domestic stock, international stock, taxable bond, and municipal bond funds. Standard errors are in the parentheses. The identifying restriction is that the sum of the performance coefficients is equal to one. The dependent variable is fund's annualized net relative flow. The control variables include a constant, logs of fund size and age, fund expense ratio and risk, log of number of funds in the family, logs of fund family size and age, objective and style category flows. Both the performance-unrelated flows and flow-performance sensitivity are linear functions of the control variables. The sample period is January 1993 - March 1999 for stock funds and January 1994 - March 1999 for bond funds. Since the coefficients on the control variables stay approximately the same as in the previous models (see Table 4.3), they are omitted from the table.

	Dom. stock		Int. stock		Tax. bond		Mun. bond	
Relative performance measures								
RP^{obj}	0.30	(0.04)	0.59	(0.07)	0.11	(0.04)	0.11	(0.24)
RP^{style}	0.14	(0.03)	0.08	(0.07)	0.12	(0.03)	0.10	(0.07)
RP^{class}	0.56	(0.05)	0.34	(0.11)	0.77	(0.05)	0.78	(0.28)
Category dummies								
D^{star}	0.39	(0.04)	0.01	(0.08)	0.30	(0.07)	1.51	(0.28)
D^{obj}	-0.03	(0.02)	-0.08	(0.07)	-0.14	(0.06)	0.14	(0.06)
D^{style}	-0.04	(0.02)	0.01	(0.07)	0.20	(0.08)	0.44	(0.08)
D^{class}	0.07	(0.01)	0.04	(0.05)	0.02	(0.02)	-0.09	(0.08)

4.8 Conclusion

Classification systems are designed to facilitate the performance evaluation of mutual funds by their investors. A classification scheme divides the fund universe into a number of categories that group funds with a similar investment philosophy. This allows investors to compare fund's performance to performance of its peers, the other funds in the category. However, there is no ideal classification scheme, which would categorize funds in truly homogenous groups. As noted by Brown and Goetzmann (1997), "there is a fundamental question whether any classification system (which, after all, is only a multinomial statistic) is sufficient to characterize differences in fund management" (p. 375). As a result, there exist several classification schemes, which categorize funds, e.g., according to their stated objective or actual investment style.

This chapter documents that performance rankings based on different classification systems are an important determinant of fund flows. The fund ranking within the asset class appears to be the most important relative performance measure for private investors of domestic stock and especially for investors of taxable bond and municipal bond funds. Less than half of the total impact of relative performance on flows to domestic stock funds is due to the stated objective and Morningstar style rankings (28% and 14%, respectively). However, institutional investors of domestic stock funds pay hardly any attention to the fund asset class rankings, using only the Morningstar style and stated objective classification systems to evaluate fund performance. In the taxable bond and municipal bond classes, the stated objective and Morningstar style rankings have a marginal impact on flows. Only in the international stock class, fund flows are most sensitive to the stated objective rankings. In a joint model of ordinal and cardinal performance measures, the impact of total return on fund flows never exceeds the combined impact of performance rankings. This implies that investors (especially those of stock funds) prefer to select funds that not only have a good performance but also outperform the others. In addition, we find that the presence of a star fund is typically beneficial for flows to the other funds in the family. Only in case of taxable bond funds, top performance of a star fund "cannibalizes" flows to funds with the same stated objective as the star fund.

The observed hierarchy of the classification schemes with respect to their importance for investors has clear implications for mutual funds and their regulators. Managers of domestic stock as well as taxable and municipal bond funds have a strong incentive to maximize their relative performance within the respective asset class, since it is used as a peer group for performance evaluation by most of their investors. However, maximizing the asset class rankings may not be consistent with the interests of fund shareholders, who are interested in maximizing the fund's risk-adjusted performance. This divergence of interests may be especially large in the domestic stock class, in which there are large differences between the risk profiles of different types of funds (say, growth funds and income funds). In particular, funds with more conservative investment styles have an incentive to take more risk to maximize the chance of outperforming their competitors, which have a riskier investment approach. Thus, classification systems on the one hand facilitate fund performance evaluation, while on the other hand they may create adverse incentives for fund managers.

Appendix 4.A Description of the Morningstar equity and fixed-income style boxes

In 1992, Morningstar introduced a style box scheme for classifying the mutual funds on the basis of their actual investment style rather than declared objective. Using the latest data on fund portfolio composition, Morningstar assigns equity and fixed-income styles to a fund, provided that sufficiently large part of the portfolio is invested in stocks and bonds, respectively. A style box is a three-by-three matrix based on two classification criterions: market capitalization and book value for equity funds; credit quality and duration for bond funds. An equity fund is classified *large*, *medium* or *small*, if the weighted-average market capitalization of the middle size quintile of its stocks falls into top 5%, the next 15% or the remaining 80% of the 5000 largest US companies, respectively. Similar procedure is applied to determine whether value or growth stocks prevail in fund's portfolio. Each fund is assigned price-to-earnings (P/E) and price-to-book (P/B) scores computed as weighted averages of P/E and P/B ratios in the respective middle quintiles. The fund is considered *growth*, *value* or *blend*, if the sum of P/E and P/B scores exceeds 2.25, is below 1.75 or falls between 1.75 and 2.25, respectively. The combination of the two criterions yields nine style categories: large value, large blend, large growth, medium value, medium blend, medium growth, small value, small blend, and small growth. The first criterion for a fixed-income style box is credit quality of the bonds in the fund portfolio. A fund is classified as *high* or *low* quality, if it has an average credit rating of AAA or AA, or lower than BBB, respectively. *Medium* quality funds fall between these two extremes. The second criterion is the interest rate sensitivity of the fund portfolio, measured as the average duration of bonds in the portfolio. Funds with an average duration less than 3.5 years, between 3.5 and 6 years, and longer than 6 years are considered as *short*, *intermediate*, and *long*, respectively. Thus, a fixed-income style box comprises high short, high intermediate, high long, medium short, medium intermediate, medium long, low short, low intermediate, and low long styles.

See Morningstar's website (www.morningstar.com) for a more detailed definition of the styles.

Chapter 5

Yet another look at tests of risk taking by mutual fund managers

5.1 Introduction

During the last two decades, the mutual fund industry experienced tremendous growth both in number of funds and amount of assets under management. It is not surprising that this industry attracts a lot of attention of the regulatory agencies that would like to ensure that fund managers select investment strategies that are optimal from the investors' point of view. The joint occurrence of two well-established facts in the mutual fund industry may lead to an agency conflict between mutual fund managers and mutual fund shareholders. First, managers' compensation is typically based on a percentage of the fund's net assets (see, e.g., Khorana, 1996). Second, the top-performing funds receive the bulk of new cash inflows, while bad performance does not lead to significant outflows (see, e.g., Sirri and Tufano, 1998). Together with the observation that at least some investors look at calendar year fund performance for their investment decisions (see, Chevalier and Ellison, 1997, p. 1183), these effects suggest that mutual fund managers participate in annual tournaments where they compete for the top rankings. This leads to the conjecture that funds performing badly during the first part of the year have an incentive to increase risk in the second part of the year in order to try to catch up with mid-year winners at the end of the year. This conjecture is called the *tournament hypothesis*. Chen and Pennacchi (1999) provide a theoretical model for risk-taking assuming that fund managers are evaluated with respect to an exogenous benchmark index. However, they show that, in their model, poor performing funds do not necessarily increase the volatility of their fund's returns.

A number of studies verifies the tournament hypothesis from an empirical point of view. Brown, Harlow, and Starks (1996) find evidence supporting the tournament hypothesis using a contingency table methodology applied to monthly data of 334 growth funds over the period 1976-1991. This technique compares volatility changes from the first to the second semester with mid-year performance. Koski and Pontiff (1999) use regression analysis and find a negative relation between interim performance and subsequent change in risk, in line with the tournament hypothesis. Koski and Pontiff (1999) use 798 domestic equity funds from 1992-1994. Finally, Chevalier and Ellison (1997), using 398 growth and growth-and-income funds from 1982-1992, obtain different regression results depending on whether fund risk is measured on the basis of fund portfolio holdings or monthly fund returns.

While previous studies, using monthly fund returns, have found strong evidence in favor of the tournament hypothesis over the period studied, Busse (2001) finds no such evidence using the contingency table methodology applied to daily data of US based funds over the 1985-1995 period. This appears surprising since daily data provide, in principle, much more precise volatility estimates and hence tests based on daily data can be expected to be more powerful in detecting evidence in favor of the tournament hypothesis than tests based on monthly data. Busse (2001) offers two explanations for this seeming paradox. First, Busse (2001) argues that biases in monthly volatility estimates due to autocorrelation in daily fund returns may adversely affect the tests based on monthly data. Such autocorrelation could be caused by mutual fund managers loading on small stocks that exhibit first-order autocorrelation due to non-synchronous trading effects (compare, e.g., Boudoukh, Richardson, and Whitelaw, 1994). Secondly, Busse (2001) notes that standard statistical tests used so far in the literature rely on the contestable assumption that mutual fund returns are cross-sectionally independent.

In this chapter, we analyze both the impact of autocorrelation and cross-correlation on test of the tournament hypothesis from an analytical point of view. First, we calculate the biases arising in volatility estimates based on monthly or daily return data due to first-order autocorrelation in the returns. More precisely, we express the ratio of the volatility over the second part of the year with respect to that over the first part of the year (henceforth, SDR or "standard deviation ratio") in terms of the autocorrelation of daily fund returns. We do this both for volatilities estimated using daily data (giving what we will henceforth call the daily SDR) and for those based on monthly data (henceforth, monthly SDR). These results show that, in line with Busse's (2001) claim, monthly SDR's are indeed (in absolute terms) more sensitive to changes in the autocorrelation pattern than daily SDR's. However, the smaller absolute bias in daily SDR's has a *larger* effect on the distribution of the tests of the tournament hypothesis, since volatility estimates based on daily data are more precise compared to volatility estimates based on monthly data. Therefore, at the end of the day, tests of the tournament hypothesis based on monthly data are more robust to autocorrelation in daily fund returns than tests based on daily data[1]. Thus, if the autocorrelation effects are such that they adversely affect the tournament tests based on monthly data, they would certainly affect the tests based on daily data. Our calculations also rationalize some of the differences in the autocorrelation of daily returns that Busse (2001) reports for funds whose monthly volatility increases/decreases during the second half of the year.

While, in this chapter, we only consider contingency table tests as in Brown, Harlow, and Starks (1996), Busse (2001) also reports empirical evidence concerning the tournament hypothesis using regression techniques similar to those of Koski and Pontiff (1999). The reader will easily convince herself that our arguments extend to those techniques. This is because our discussion concerning both the bias in volatility estimates and the required form of (in)dependence in mutual fund returns are independent of the actual test employed, be it a contingency table test or a regression based tests. Busse (2001) also presents tests of the tournament hypothesis based on ratios of the residual volatility of fund returns in an MA(1) model. The conclusions from these tests for the 1985-1995 sample period are the same as those based on the total return volatility (see Table 2 and Table 3 of Busse, 2001). These test results are not corrected for cross-correlation effects in fund returns.

As mentioned before, a second contribution of the present chapter is the derivation of explicit conditions for the validity of the tournament hypothesis tests. We show that these tests implicitly assume that fund returns follow a factor structure with *uncorrelated* idiosyncratic errors across funds. This is what Chamberlain and Rothchild (1983) have named a *strict factor structure*. If this hypothesis is not satisfied, size corrected p-values can be obtained using simulation or bootstrap techniques. When using these size

[1] Formally, the non-centrality parameter in the χ^2-distribution arising from autocorrelation effects, which is given by the squared bias over the sampling variance, is smaller if one uses monthly data.

corrected p-values, the evidence in favor of the tournament hypothesis based on monthly fund returns disappears. Busse (2001) also reports size corrected test results based on daily fund returns, finding again no evidence in favor of the tournament hypothesis.

The structure of this chapter is as follows. In Section 5.2, we present the analytical results concerning the bias arising in volatility estimates based on daily and monthly data and show that tests of the tournament hypothesis based on daily data are more severely affected by autocorrelation effects than tests based on monthly data. Using these calculations, we also rationalize part of the results reported in Busse (2001) about the relation between monthly volatility changes and changes in the autocorrelation pattern of fund returns. Section 5.3 discusses the conditions under which the contingency table tests have the appropriate size and gives an illustrative empirical example. Section 5.4 concludes and Appendix 5.A gathers some proofs.

5.2 Effect of autocorrelated returns on SDR's

Following Busse (2001), we consider first-order moving average (MA(1)) specifications for daily fund returns. More precisely, using the same notation as Busse (2001), we have, for fund p and day d,

$$
\begin{aligned}
r_{p1d} &= \mu_{p1} + \theta_{p1}\varepsilon_{p1,d-1} + \varepsilon_{p1d}, \quad d = 1, \ldots, D, &(5.1)\\
r_{p2d} &= \mu_{p2} + \theta_{p2}\varepsilon_{p2,d-1} + \varepsilon_{p2d}, \quad d = D+1, \ldots, D_y, &(5.2)
\end{aligned}
$$

where $d = 1, \ldots, D$ refers to the first part of the year (subindex "1") and $d = D + 1, \ldots, D_y$ refers to the second half of the year (subindex "2"). In Appendix 5.A, we calculate the biases in the daily and monthly SDR's as a function of the autocorrelation coefficients θ_{p1} and θ_{p2}. These standard deviation ratios are crucial in the contingency table tests for the tournament hypothesis that we shortly describe now. The daily SDR for fund p is given by

$$
SDR_p = \sqrt{\frac{\frac{1}{D_y - D - 1}\sum_{d=D+1}^{D_y}\left(r_{p2d} - \bar{r}_{p2}\right)^2}{\frac{1}{D-1}\sum_{d=1}^{D}\left(r_{p1d} - \bar{r}_{p1}\right)^2}}, \tag{5.3}
$$

where $\bar{r}_{p1}$ denotes the average return over the first part of the year and $\bar{r}_{p2}$ that over the second part of the year. Let $Med(\bar{r}_{p1})$ denote the median average fund return over the first part of the year and let $Med(SDR_p)$ denote the median SDR. The contingency table test statistic for the tournament hypothesis given in Brown, Harlow, and Starks (1996) can now be written as

$$Q = \left(\frac{\text{number of funds with } \bar{r}_{p1} < Med(\bar{r}_{p1}) \text{ and } SDR_p > Med(SDR_p)}{\text{total number of funds}} - 0.25 \right)^2 . \tag{5.4}$$

Under the null hypothesis that past returns and subsequent risk-taking are independent, and that returns are distributed independently over funds, Q follows asymptotically a χ^2 distribution with one degree of freedom. The corresponding critical values are routinely used in many empirical studies. In Section 5.3 and Appendix 5.A the distribution of the test statistic Q is derived under more realistic assumptions.

From the results in the appendix, we find in first-order approximation and under the null of constant idiosyncratic volatility during the year (i.e., $\sigma_{\varepsilon p1} = \sigma_{\varepsilon p2}$),

$$\text{Monthly SDR bias} \approx \theta_{p2} - \theta_{p1}, \tag{5.5}$$

$$\text{Daily SDR bias} \approx \bar{\theta}_p(\theta_{p2} - \theta_{p1}), \tag{5.6}$$

where $\bar{\theta}_p$ denotes the average MA(1) coefficient for fund p, i.e., $\bar{\theta}_p = (\theta_{p1} + \theta_{p2})/2$.

Busse (2001) claims that first-order autocorrelation in daily fund returns biases monthly volatility estimates. The effect of the autocorrelation on the monthly SDR is apparent from equation (5.5), and substantiates Busse's (2001) claim. If, due to external reasons, fund managers load more on less liquid stocks during the second part of the year, they may increase first-order autocorrelation in daily fund returns and thereby influence the monthly SDR. However, also the daily SDR's are affected by the changing autocorrelation, albeit to a lesser extent since $\bar{\theta}$ is (see Table 5 of Busse, 2001) at most 0.2. The relevant question is what the effect of changing autocorrelation is on the contingency table tests. This effect is measured by the non-centrality parameter in the χ^2_1-distribution caused by the daily autocorrelation. This non-centrality parameter is given by the squared bias over the estimation variance (see, e.g., Godfrey, 1991, p. 18). This quantity can be calculated using the relative efficiency of both SDR estimates as given in Busse (2001). Given the reported daily MA(1) coefficients in Busse (2001) of about 0.17[2], the squared monthly SDR bias is about $(1/0.17)^2 = 34.60$ times that of the squared daily SDR bias. The relative efficiency of daily SDR's with respect to monthly

[2] This number is calculated as the MA(1) coefficient that induces a first-order autocorrelation which equals the average autocorrelation reported in Panel B of Table 5 in Busse (2001).

SDR's is given in Busse (2001) as about 47.3[3]. Comparing the squared biases to the relative efficiency of the volatility estimates, we see that the contingency table test based on monthly data is more robust to changes in the first-order autocorrelation of daily returns than tests based on daily data. The autocorrelation effect alone can thus not account for the difference in empirical evidence concerning the tournament hypothesis. If changes in the first-order autocorrelation of daily fund returns, possibly generated by changes in the fund's loading on small stocks, affect the contingency table tests based on monthly fund returns, they will also affect the tests based on daily fund returns and in the same direction. However, Table 2 in Busse (2001) shows that, for some specifications, the monthly tests seem to confirm the tournament hypothesis, while the daily tests do not. In all cases, the direction of the rejection of the monthly and daily tests is opposite.

The analytical relation between monthly SDR's and the MA(1) coefficients of first and second semester returns in (5.10) also explains some of the findings in Section IV.C, and in particular in Table 5, of Busse (2001). Busse (2001) reports that the funds classified as high monthly SDR have smaller average January-June MA(1) coefficients than the corresponding low monthly SDR funds. Also, funds classified as high monthly SDR have larger average increases in autocorrelation from the beginning to the end of the year. These reported results are in line with equation (5.5). When selecting high monthly SDR funds, one selects those funds that have a relatively large MA(1) coefficient during the second half of the year (θ_{p2}) as compared to that during the first part of the year (θ_{p1}). The fact that all funds on average show an increase in autocorrelation from the first to the second part of the year, has no repercussions for this argument.

Clearly, the analysis in this section refers to the marginal effect of autocorrelation, not taking into account a possible interaction with the cross-sectional dependence effect which is also mentioned in Busse (2001). In the next section, we study the effect of cross-sectional dependence in more detail by providing explicit conditions under which the tournament tests employed so far in the literature have the correct size. Again, for expository reasons, we focus on the contingence table test of Brown, Harlow, and Starks (1996), but the results readily extend to regression based tests.

[3]This number is reported by Busse at the end of Section III.B. It is calculated neglecting serial correlation in daily fund returns.

5.3 Contingency table test and strict factor structure

In this section, we take a closer look at the conditions needed for the contingency table test statistic (5.4) to have a χ_1^2 limiting null distribution. Busse (2001) notes that cross-correlation in fund returns violates the independence assumption used in deriving the χ^2 tests for equal cell frequencies. In the appendix, we show that actual independence of fund returns is not necessary for the contingency table test statistic to have a χ^2 limiting null distribution. As long as fund returns are generated from a factor model in which the idiosyncratic returns are uncorrelated across funds, χ^2 critical values may be used to obtain a test with the required size. Such a factor structure is what Chamberlain and Rothchild (1983) have named a strict factor structure.

Busse (2001) adopts a bootstrap technique to simulate p-values that are robust to possible cross-correlation of fund returns. He reports that, both for daily and monthly fund returns and using these robust p-values, no evidence in favor of the tournament hypothesis is found in the sample under consideration. If the cross-correlation is taken into account, neither daily nor monthly fund returns point in the direction of strategic risk taking by mutual fund managers. To illustrate this point and to make the chapter self-contained, we perform the contingency table test for strategic risk-taking on monthly data for 19 annual tournaments. Similar tables are given in Brown, Harlow, and Starks (1996) and Busse (2001). Our dataset consists of 811 US growth funds from the Morningstar Mutual Fund Database. The time span is from 1976 to 1994 including the 1989-1994 time period for which our data are free from survivorship bias. In line with Busse (2001), we split years in two equal periods of six months when calculating monthly SDR's.

Table 5.1 reports the results. Using χ_1^2-based p-values, we find significant results in almost half of the 19 annual tournaments. However, again for about half of the years for which results appear significant, the results are opposite to the tournament hypothesis, i.e., losing funds are found to (relatively) reduce risk-taking over the second part of the year. In order to accommodate possibly cross-correlated idiosyncratic monthly mutual fund returns, we determine the distribution of the tournament tests, under the null of no strategic risk-taking, using simulation. To explain the procedure in more detail, for each month, we simulate the vector of fund returns from a multivariate normal distribution

Table 5.1: **Results of the contingency table approach for the 19 annual tournaments in 1976-1994**

The Low-High column gives the percentage of funds with both a total return over the first six months below median and a risk adjustment ratio (SDR) above median. The χ^2-statistic tests the null hypothesis that population percentages are equal to 25%. Column five presents the p-values of the χ^2-statistic based on the χ^2_1 distribution. The last column reports simulated p-values for the χ^2-statistic. See main text for details.

Year	# funds	Low-High frequency	χ^2-statistic	p-value (χ^2_1)	p-value (simulated)
1976	119	25.21	0.01	0.9270	0.9409
1977	123	16.26	15.03	0.0001	0.1393
1978	128	14.06	24.50	0.0000	0.0697
1979	131	26.34	0.37	0.5408	0.7637
1980	132	27.27	1.09	0.2963	0.6629
1981	139	21.58	2.60	0.1071	0.5520
1982	144	23.61	0.44	0.5050	0.7146
1983	160	17.19	15.63	0.0001	0.0939
1984	181	24.03	0.27	0.6029	0.8304
1985	203	24.88	0.00	0.9440	0.9543
1986	235	23.40	0.96	0.3278	0.7357
1987	271	31.00	15.59	0.0001	0.2450
1988	315	29.68	11.05	0.0009	0.3986
1989	336	27.68	3.86	0.0495	0.5732
1990	357	29.69	12.57	0.0004	0.2968
1991	392	34.82	60.50	0.0000	0.0386
1992	425	24.94	0.00	0.9613	0.9816
1993	520	18.75	32.50	0.0000	0.1852
1994	635	26.30	1.71	0.1903	0.7638

with a mean vector and variance matrix that are estimated from the observed monthly fund returns. In these simulated fund returns there is no tournament effect. For each null simulation, we calculate the realization of the contingency table test statistic (5.4). This is replicated 10,000 times from which the simulated p-values are obtained.

This way of simulating robust p-values has a particularly nice invariance property. If actual fund returns would be generated from a factor model, this would not invalidate our way of simulating fund returns (which is, in fact, based on a zero-factor model). To see why this is true, suppose that a factor model had been estimated and that fund returns were simulated using the estimated factor loadings, the observed factor values, and the variance matrix of idiosyncratic fund returns. In the end, the simulated fund returns would then again be normally distributed with exactly the same mean vector and variance matrix as above. This follows immediately from the standard orthogonality of the regression decomposition. Our zero-factor simulation hence generates fund returns that are distributionally equal to those generated from any other factor model.

Our simulations assume normality of monthly fund returns. Clearly, this normality assumption is innocuous, if sufficient regularity conditions are satisfied for a central limit theorem to hold true. It is important that the simulation setup allows (idiosyncratic) fund returns to be correlated across funds. Busse (2001) uses bootstrapped critical values. The advantage of such an approach is that it does not rely on any normality assumptions, which may be relevant when using daily data. The disadvantage is that it is computationally somewhat more intensive. Moreover, as Busse (2001) reports, p-values based on a Monte Carlo approach assuming normality do not produce materially different results from p-values based on a bootstrap approach.

The last column of Table 5.1 reports the cross-correlation robust p-values. In line with Busse (2001), all evidence in favor of the tournament hypothesis disappears once the cross-correlation is accounted for. Thus, when using corrected p-values, monthly and daily fund returns lead to the same conclusion, where, given the increased efficiency, tests based on daily fund returns can be more powerful in detecting evidence in favor of the tournament hypothesis.

5.4 Conclusion

The present chapter confirms the conclusion in Busse (2001) that, for US equity funds over the sample periods studied so far, there is little empirical evidence in favor of the tournament hypothesis for mutual fund managers. We study in detail the reasons for the difference in empirical results of the tournament test based on daily and monthly data, from an analytical point of view. We argue that the source of spurious evidence found in the past is not so much neglected temporal correlation in returns, but neglected cross-correlation between idiosyncratic fund returns. Autocorrelation in daily fund returns indeed biases both monthly and daily SDR's, but tests based on monthly SDR's prove to be more robust to these effects than tests based on daily SDR's. Thus, spurious (due to autocorrelation effects) evidence in favor of the tournament hypothesis based on monthly returns, would, ceteris paribus, certainly show up in tournament tests based on daily returns for the same sample period.

On the other hand, neglecting cross-correlation in fund returns may lead (as already noted in Busse, 2001) to spurious inference. We show that cross-correlated fund returns do not necessarily invalidate the tournaments tests used so far in the literature, as long as the idiosyncratic fund returns in some factor model are uncorrelated across funds. When cross-correlation is accounted for, all empirical evidence in favor of the tournament hypothesis, based on methodology used in previous studies and our 1976-1994 sample of US growth funds, disappears.

Appendix 5.A Biases in daily and monthly SDR's

We take continuously compounded daily returns r_{pjd} as given in (5.1)-(5.2). If D_m denotes the number of days in a month, the monthly returns, for month m, are given as

$$r_{pjm} = \sum_{d=1}^{D_m} r_{pjd}, \quad j = 1, 2. \tag{5.7}$$

It is well-known that the stationary variance of daily fund returns during the first half of the year, for fund p, is $(1+\theta_{p1}^2)\sigma_{\varepsilon p1}^2$, where $\sigma_{\varepsilon p1}$ denotes the idiosyncratic volatility of fund p, i.e., $\sigma_{\varepsilon p1}^2 = \text{var}\{\varepsilon_{p1}\}$. This immediately implies

$$\text{Daily SDR} = \sqrt{\frac{\text{var}\{r_{p2d}\}}{\text{var}\{r_{p1d}\}}} = \sqrt{\frac{1+\theta_{p2}^2}{1+\theta_{p1}^2}}\frac{\sigma_{\varepsilon p2}}{\sigma_{\varepsilon p1}} \approx \frac{1+\frac{1}{2}\theta_{p2}^2}{1+\frac{1}{2}\theta_{p1}^2}\frac{\sigma_{\varepsilon p2}}{\sigma_{\varepsilon p1}}, \tag{5.8}$$

where the latter approximation is immediate from $\sqrt{1+x^2} \approx 1+\frac{1}{2}x^2$, for small x. Note that Busse (2001) reports MA(1) coefficients θ in the interval from 0.0 to 0.2, so that the squared autocorrelation coefficient is at most 0.04.

It is somewhat more complicated to calculate the monthly SDR. From the autocorrelation function of an MA(1) process, we obtain

$$\begin{aligned} \text{var}\{r_{p1m}\} &= D_m \text{var}\{r_{p1d}\} + 2(D_m - 1)\text{cov}\{r_{p1d}, r_{p1,d-1}\} \\ &= D_m(1+\theta_{p1}^2)\sigma_{\varepsilon p1}^2 + 2(D_m-1)\theta_{p1}\sigma_{\varepsilon p1}^2, \\ &= D_m(1+\theta_{p1})^2\sigma_{\varepsilon p1}^2 - 2\theta_{p1}\sigma_{\varepsilon p1}^2, \end{aligned} \tag{5.9}$$

and a similar relation for the second half of the year. Therefore, the monthly SDR is given by

$$\text{Monthly SDR} = \sqrt{\frac{\text{var}\{r_{p2m}\}}{\text{var}\{r_{p1m}\}}} = \sqrt{\frac{D_m(1+\theta_{p2})^2 - 2\theta_{p2}}{D_m(1+\theta_{p1})^2 - 2\theta_{p1}}}\frac{\sigma_{\varepsilon p2}}{\sigma_{\varepsilon p1}} \approx \frac{1+\theta_{p2}}{1+\theta_{p1}}\frac{\sigma_{\varepsilon p2}}{\sigma_{\varepsilon p1}}, \tag{5.10}$$

an approximation based on the fact that D_m is much larger than the MA(1) coefficient.

Under the null of constant idiosyncratic volatility during the year (i.e., $\sigma_{\varepsilon p1} = \sigma_{\varepsilon p2}$),

$$\text{Monthly SDR bias} = \text{Monthly SDR} - 1 \approx \theta_{p2} - \theta_{p1}, \tag{5.11}$$

$$\text{Daily SDR bias} = \text{Daily SDR} - 1 \approx \frac{1}{2}(\theta_{p2}^2 - \theta_{p1}^2) = \frac{1}{2}(\theta_{p1}+\theta_{p2})(\theta_{p2}-\theta_{p1}), \tag{5.12}$$

since for small x and y we have $\frac{1+x}{1+y} \approx x - y$.

Appendix 5.B Limiting distribution of the contingency table test

In this appendix, we derive the limiting distribution of the contingency table test mentioned in the main text assuming that monthly mutual fund returns are generated from a strict factor model. Clearly, the same results would hold true for any other data frequency. We assume for the moment that, for fund p in month m,

$$r_{pm} = \alpha_p + \beta_p^T F_m + \varepsilon_{pm}, \tag{5.13}$$

where F_m denotes the vector of factors and where the idiosyncratic errors ε_{pm} are independently $N(0, \sigma_p^2)$ distributed[4]. Define the sample average and the sample volatility of fund p's returns over some period j consisting of k months as $\bar{r}_{pj}$ and $\hat{\sigma}_{pj}$, respectively:

$$\begin{aligned} \bar{r}_{pj} &= \frac{1}{k} \sum_{m=1+k(j-1)}^{kj} r_{pm}, \\ \hat{\sigma}_{pj} &= \sqrt{\frac{1}{k} \sum_{m=1+k(j-1)}^{kj} (r_{pm} - \hat{\mu}_p^{(j)})^2}. \end{aligned}$$

Let $\mathcal{F}$ denote the information in the factors over the complete observational period, i.e. $\mathcal{F} = \sigma(F_1, F_2, \ldots)$. Now, conditionally on $\mathcal{F}$ and under the null hypothesis, the statistics $\bar{r}_{p1}$, $\bar{r}_{p2}$, $\hat{\sigma}_{p1}$, and $\hat{\sigma}_{p2}$ are independently distributed. The independence of the risk-adjustment-ratio $\hat{\sigma}_{p2}/\hat{\sigma}_{p1}$ and the first semester average return $\bar{r}_{p1}$, implies that (conditionally on $\mathcal{F}$ and under the null) the standard χ^2-test statistic Q for independence of risk-adjustment-ratios and first semester returns (5.4) follows, asymptotically, a χ_1^2 distribution. Formally, under the null hypothesis,

$$\mathcal{L}(Q|\mathcal{F}) \longrightarrow \chi_1^2.$$

For regression tests, the independence of the idiosyncratic errors ε_{pm} across funds, would guarantee the validity of the standard t-test by the same arguments.

In case the idiosyncratic errors ε_{pm} are correlated across funds, the arguments above no longer hold, even asymptotically. In that case, the number of unbounded eigenvalues of the variance of fund returns is generally infinite and limiting results can no longer be established analytically in general.

[4]Clearly, normality is, asymptotically, irrelevant for the main results in this appendix as long as variances exist since then one may resort to a central limit theorem argument.

Chapter 6

Mutual fund tournament: Risk taking incentives induced by ranking objectives

6.1 Introduction

When choosing between mutual funds, investors take into account many considerations such as fund performance, reputation, fee structure, the diversity and size of the fund's family, etc. (see, e.g., Chevalier and Ellison, 1997, and Sirri and Tufano, 1998). Naturally, a rational investor will select a fund, which offers the best combination of the relevant factors. Since fund performance seems to be the most important selection criterion for consumers (see Capon, Fitzsimons, and Prince, 1996), they typically choose funds that have high raw or risk-adjusted performance relative to their peer group (see, e.g., Ippolito, 1992, Chevalier and Ellison, 1997, Lettau, 1997, and Sirri and Tuffano, 1998). The information about fund performance rankings is regularly published in the financial media (the Wallstreet Journal, Business Week, Money, etc.) and is often referred to in funds' advertisements. The importance of rankings in describing fund performance is illustrated by Gould (1998):

> *"Bartlett Europe has returned an annual average of 27.2 percent for the three years through Dec. 4, ranking first among the 46 European stock funds tracked by Morningstar Inc."*

The academic literature also points out the importance of fund performance rankings for investors, documenting that rankings may have higher impact on fund flows than returns (see, e.g., Patel, Zeckhauser and Hendricks, 1994, Massa, 1997, and Chapter 4). Such investors' behavior induces ranking-based objectives for fund managers, since their compensation is typically based on a percentage of the fund's assets (see Khorana, 1996).

The goal of this chapter is to investigate how ranking objectives influence managers' investment strategies, and test empirically some predictions of the model. In the first part of the chapter, we develop a model in which, during two investment periods, two risk-neutral managers compete for future money flows and observe their interim relative performance. We show that in the first period, managers choose the same risk level but do not maximize their expected return. In the second period, the interim loser *(i)* increases risk with respect to the first period while the interim winner decreases risk, *(ii)* the difference in risk undertaken is increasing with the difference in interim performances and *(iii)* the interim loser may act more in the interest of investors (i.e., choose a strategy with a higher expected return) than the interim winner.[1]

In the second part of the chapter, we apply a new methodology to empirically test some predictions of the model in a sample of US diversified equity funds in 1980-1998. We find evidence that fund choice of systematic risk in the last quarter of the year is negatively related to the category-relative performance over the first three quarters of the year, which is consistent with the model. In contrast to the previous studies (see, e.g., Brown, Harlow and Starks, 1996, and Koski and Pontiff, 1999), our statistical tests take into account the presence of the cross-correlation in fund returns highlighted by Busse (2001).

The organization of this chapter is as follows. Section 6.2 reviews the related literature, Section 6.3 presents the model, Section 6.4 derives the equilibrium, Section 6.5 considers the case with several competing funds, Section 6.6 presents the empirical results, and Section 6.7 concludes.

[1] Using simulations, we provide evidence that manager's choice of risk in the second period is negatively related to his relative performance in the first period also in the case with more than two competing funds.

6.2 Related literature

A growing body of literature studies the mutual fund tournament both theoretically and empirically. Closely related theoretical papers studying relative performance evaluation in financial markets are those of Huddart (1999), Hvide (1999), Palomino (2002), and Taylor (2000). In this type of the models, a manager's payoff depends not only on his own strategy, but also on the other managers' strategies. In this respect, these models and our model are different from those analyzing the behavior of a manager evaluated against an exogenous benchmark (see, e.g., Grinblatt and Titman, 1989, Admati and Pfleiderer, 1996, Carpenter, 2000).

Hvide (1999) and Palomino (2002) study the consequences of relative performance objective in the context of a single investment decision. Hvide shows that in a situation with moral hazard on both effort and risk, standard tournament rewards induce excessive risk and lack of effort. Palomino (2002) assumes that managers with different levels of information compete in oligopolistic markets and aim at maximizing their relative performance against the average performance in their category. He shows that despite the objective function being linear in performances, managers have incentives to choose overly-risky strategies. Furthermore, relative performance objectives always lead to under-acquisition of information. Huddart (1999) considers a two-period model in which interim performances are observable. He shows that asset-based compensation schemes generate incentives for managers to invest in overly-risky portfolio in the first period, and that performance fees align managers' incentives with those of investors.

Our theoretical results should be compared with those of Cabral (1997) on the choice of R&D projects. Cabral considers an infinite-period race between two firms that choose between low variance projects (low gains with high probability) and high variance projects (large gains with low probability). If the two firms choose a project of the same type then outcomes are perfectly correlated. Cabral shows that in equilibrium, both firms choose overly risky R&D strategies. There are three main differences between Cabral's model and ours. First, in Cabral's model, players have an infinite horizon. It follows that strategy choices are not influenced by an "end of the game" effect. Second, players receive a payoff in every period. This is equivalent to assuming observable interim performance. Conversely, in our model, players face an end of the game and only receive a payoff at the end of the game. Finally, in Cabral's R&D race, projects' payoffs are different only in case of success. If projects fail, the costs faced by firms are independent of the projects chosen. This implies that an intermediate loser only catches up with the leader if a good outcome is realized. The situation is different in the mutual fund tournament. An intermediate loser has two ways of catching up with the winner: by winning more in case of good outcomes or by losing less is case of bad outcomes.

The consequences of dynamic incentives and relative performance evaluation have also been studied by Meyer and Vickers (1997). They show that in a dynamic principal-agent relationship, relative performance evaluation can be either welfare increasing or decreasing. The reason is that in a dynamic setting, there may be both explicit and implicit incentives and better information may decrease implicit incentives. Our model is different from that of Meyer and Vickers in two ways. First, in their model, intermediate performance is observable. Second, in our model, portfolio decisions are costless, i.e., they do not require any effort from fund managers. This is different from standard principal-agent models in which agents' output results from costly effort.

Another strand of the literature conducts empirical analysis of fund managers' strategic behavior, focusing on the impact of past performance on funds' risk taking decisions. Several studies test the so-called tournament hypothesis that funds underperforming after the first part of the year increase risk in the second part of the year, trying to catch up with interim winners at the end of the year. Applying a contingency table methodology to the sample of US growth funds in 1976-1991, Brown, Harlow, and Starks (1996) find that interim losers (defined as funds below the median category return over the first part of the year) increase risk towards the end of the year relative to interim winners. Using a sample of US domestic equity funds in 1992-1994, Koski and Pontiff (1999) apply regression methodology and find a negative relationship between fund return over the first semester and the change in total, systematic, and unsystematic risk between the first and second semesters. Chevalier and Ellison (1997) use a different approach, measuring fund risk on the basis of the fund's portfolio holdings. They also find a negative relationship between fund return over the first nine months of the year and the change in fund risk between September and December, using a sample of growth and growth-and-income funds in 1982-1992. However, Busse (2001) argues that these results should be taken with caution. He finds no evidence in favor of the tournament hypothesis, applying either the contingency table or the regression methodology to daily returns of US domestic equity funds in 1985-1995. He explains this divergence in the results by the presence of the auto- and cross-correlation in fund returns, which was not accounted for in the standard statistical tests used in the previous studies.

6.3 Presentation of the model

There are two periods, 1 and 2, and two risk-neutral money managers. At the beginning of the first period, each manager has $A \geq 0$ units of money under management. At the beginning of each period, managers choose an investment strategy. There is a continuum of investment strategies and the return of each strategy is log-normally distributed. The log-return of a strategy is normally distributed with variance v and mean $m(v)$. Following Palomino and Prat (2002), we assume that the function $m(\cdot)$ is positive, twice differentiable, strictly concave with $m'''(\cdot) \geq 0$, and has a maximum at $\hat{m} = m(\hat{v})$ with $\hat{v}$ strictly positive.

A possible interpretation for the shape of $m(\cdot)$ is that there is no borrowing constraint but borrowing is increasingly costly. Therefore, there is a borrowing threshold beyond which the marginal borrowing cost exceeds the marginal expected return of investment.

Information about realized returns. After returns are realized at the end of period 1, managers observe both their performance and the performance of their opponent.

Managers' objective: Managers aim at maximizing the size of the funds under management at the end of period 2. The fund size can be increased in two ways. First, by realizing a high cumulated return over the periods 1 and 2, and second by attracting new funds.

There is a continuum of identical atomistic individual investors. On aggregate, these investors will have an amount of money $B > 0$ to invest at the end of period 2. Following empirical evidence provided by Patel, Zeckhauser and Hendricks (1994) and Massa (1997), we assume that investors put their money in the fund that has realized the higher cumulative return over periods 1 and 2. If funds perform equally well, each fund will get an amount $B/2$.

Under such an assumption, the objective of manager i is to maximize

$$C_i = \begin{cases} AR_{i,1}R_{i,2} + B & \text{if} \quad R_{i,1}R_{i,2} > R_{j,1}R_{j,2} \\ AR_{i,1}R_{i,2} + B/2 & \text{if} \quad R_{i,1}R_{i,2} = R_{j,1}R_{j,2} \\ AR_{i,1}R_{i,2} & \text{if} \quad R_{i,1}R_{i,2} < R_{j,1}R_{j,2} \end{cases} \tag{6.1}$$

with $j \neq i$, and where $R_{i,t}$ represents the gross return realized by manager i in period t.

Our model captures the following idea in a simple framework. First, investors use rankings as a rule of thumb to evaluate managers and allocate capital to funds (as empirical evidence provided by Patel, Zeckhauser and Hendricks (1994) and Massa (1997) suggests). Second, fund managers are risk-neutral agents who maximize the size of the fund they manage.

Before proceeding, several remarks should be made. First, we do not question whether fund investors are right or wrong to use rankings as a rule of thumb to evaluate fund managers. Rather, we study the *consequences* of this observed behavior.

Second, following Das and Sundaram (2002) and Palomino and Uhlig (2002), we depart from the traditional principal-agent approach to contracting in which the principal (i.e., the investor) decides the compensation contract of the agent (i.e., the fund manager). In the mutual fund industry, funds (i.e., agents) choose the type of fee they charge to investors (principals). In our model, the compensation scheme (i.e., an asset-based compensation) is given. However, analyzing a game in which compensation contracts are endogenous, Palomino and Uhlig (2002) derive conditions under which asset-based contracts are chosen by mutual fund managers in equilibrium.

Third, it is assumed that portfolios are unobservable. We believe that this assumption is realistic, since portfolio disclosures are not frequent[2] and managers window-dress their portfolio around disclosure dates in practice (see, e.g., Musto, 1999, and Carhart et al., 2002).

Also, we assume that returns realized by managers are uncorrelated. This implies that the only strategic decision of the managers is the variance of their portfolio. A more complete model would assume that a manager can also influence the covariance of returns. Such a case is considered in Appendix 6.B and it is shown that, qualitatively, the results about risk taking incentives generated by ranking objectives derived in the case of uncorrelated returns still hold in the case of correlated returns.

[2]In the United States, mutual fund portfolios have to be disclosed semiannually. However, other countries such as the Netherlands only require disclosure once a year.

Finally, it can be argued that investors evaluate managers with respect to each other only if the two managers are of different qualities. This may not be the case. It is sufficient that investors *believe* that managers are of different qualities. For example, consider the following situation. With probability 1/2, manager i is a high quality manager and with probability 1/2 he is a bad quality manager, and probabilities of being a good manager are independent across managers. Moreover, the two managers observe the realized types while investors do not. In such a situation, with probability 1/2, it is common knowledge among managers that they are of the same type. However, investors do not know whether managers are of the same type. According to investors' beliefs, with probability 1/2, there is a good and a bad manager, and they use a relative performance rule to evaluate managers.

Here, in order to concentrate on incentives generated by differences in intermediate performances, we solely study the case in which managers are of the same quality. If managers were of different qualities, incentives in period 2 would be driven by both interim performances and difference in quality.

The benchmark case. We consider as a benchmark the case in which managers maximize their expected return. In such a situation, in each period, the expected return of a manager is $m(v) + v/2$. Hence, both managers choose a risk level $v = \bar{v}$ such that $m'(\bar{v}) = -\frac{1}{2}$. The goal of our model is to show how ranking objectives alter the managers' investment strategies.

6.4 Equilibrium investment strategies

We solve the model using backward induction. Hence, we start by deriving the equilibrium of the game played by the two managers in period 2. Denote $R_{t,w}$ and $R_{t,l}$ the gross return obtained in period by t by the interim winner and loser, respectively. Let $r_{j,t} = \log(R_{j,t})$ ($j = l, w$ and $t = 1, 2$) and $\delta = r_{w,2} - r_{l,2}$. From the assumption about the distribution of returns, $r_{l,2} - r_{w,2}$ is normally distributed with mean $m(v_l) - m(v_w)$ and variance $v_l + v_w$. Hence, the objective of the interim loser is to maximize

$$H_l(v_l, v_w, \delta) = A \exp\left(m(v_l) + \frac{1}{2}v_l\right) + B\left[1 - \Phi\left(\frac{\delta + m(v_w) - m(v_l)}{(v_l + v_w)^{1/2}}\right)\right] \tag{6.2}$$

over v_l, while the objective of the interim winner is to maximize

$$H_w(v_w, v_l, \delta) = A \exp\left(m(v_w) + \frac{1}{2}v_w\right) + B\Phi\left(\frac{\delta + m(v_w) - m(v_l)}{(v_l + v_w)^{1/2}}\right) \tag{6.3}$$

over v_w, where Φ is cdf of the standard normal distribution.

A manager's objective is to maximize the size of his fund at the end of the second period. This can be achieved in two ways. First, by obtaining a high return. This is captured by the first term in the right-hand sides of (6.2) and (6.3). This provides managers with incentives to maximize their expected return (i.e., choose $v = \bar{v}$). The second way of increasing the size of the fund is by outperforming the opponent. This is captured by the second term in the right-hand sides of (6.2) and (6.3). The larger the ratio A/B, the more managers' incentives are aligned with investors' interests (i.e., the maximization of expected returns). Conversely, when the ratio A/B is small, managers' main objective is to outperform their opponent in order to receive B. To isolate incentives generated by tournament objectives, we concentrate on the case in which A is negligible with respect to B. (Technically, we assume that $A = 0$.) In such a situation, managers' only objective is to be ranked first. We have the following results.

Proposition 6.1 *Assume that managers' objective function is given by (6.1) with $A = 0$.*

(i) *If $\delta \neq 0$, then in the second period, the unique equilibrium is such that $v_w^* < \hat{v} < v_l^*$ with $|m'(v_w^*)| = |m'(v_l^*)|$. Furthermore, v_l^* and v_w^* are increasing and decreasing in δ, respectively.*

(ii) *If $\delta = 0$, then there exists a unique symmetric equilibrium in the second period: both managers choose $\hat{v}$.*

Proof: See Appendix 6.A.

Proposition 6.1 states that if the two funds have performed differently in the first period ($\delta \neq 0$), then, in the last period, the unique equilibrium is such that an interim loser takes more risk than an interim winner. Furthermore, the larger the difference in performance between the interim winner and the interim loser at the end of period 1, the larger the difference in risk undertaken in period 2. If managers have performed equally well in the first period ($\delta = 0$), they both choose a conservative strategy ($\hat{v}$) in the first period. The reason for this last result is that if manager i chooses $\hat{v}$ and manager $j \neq i$ does not, then manager i has a probability of winning the contest strictly larger than 1/2, while if manager j chooses $\hat{v}$, both managers have a probability 1/2 of winning the contest. Conversely, $\bar{v}$ is never a best reply to $\bar{v}$, the reason being that the distribution of returns is not symmetric around its mean.

By the same argument, we derive equilibrium strategies played in the first period.

Proposition 6.2 *Assume that managers' objective function is given by (6.1) with $A = 0$. There exists a unique symmetric equilibrium in the first period: both managers choose $\hat{v}$.*

Proof: See Appendix 6.A.

Proposition 6.2 implies that in the first period managers do not act in the interest of investors. The reason is that the log-normal distribution is not symmetric with respect to its mean. It follows that if one manager chooses $v = \bar{v}$, then the best reply of his opponent is not $\bar{v}$.

From Propositions 6.1 and 6.2, we deduce that when compensation is exclusively based on ranking, an interim winner locks in his gain in the second period, hence decreasing his level of risk undertaken with respect to the first period. Conversely, the interim loser increases risk with respect to the first period. Note, however, that if δ is small, we have $v_w^* < \hat{v} < v_l^* < \bar{v}$. This implies that in the second period the interim loser acts more in the interest of investors than interim winners.

6.5 More than two competing funds

So far, we have assumed that there are only two competing funds. In this section, we consider a more realistic case in which there are more than two competing funds. Denote $N > 2$ the number of competing funds and assume that fund i receives the investors' money if it has the highest return over two periods:

$$\begin{array}{ll} C_i = B & \text{if } R_{i,1}R_{i,2} > R_{j,1}R_{j,2} \ (i \neq j) \\ C_i = 0 & \text{otherwise} \end{array} \tag{6.4}$$

Let $\delta_{ij} = r_{i,1} - r_{j,1}$. The objective of fund i in period 2 is to maximize

$$\text{Prob}\left(r_{i,2} > \max_{j \neq i}(r_{j,2} - \delta_{ij})\right).$$

If fund returns are uncorrelated, this is equivalent to maximizing

$$\prod_{j \neq i} \text{Prob}\left(r_{2,i} > r_{2,j} - \delta_{ij}\right).$$

Let

$$G(v_i, v_j, \delta_{ij}) = \frac{\delta_{ij} + m(v_i) - m(v_j)}{(v_i + v_j)^{1/2}}.$$

Given that log-returns are normally distributed, the first-order condition of the maximization program of manager i ($i = 1, \ldots, N$) is

$$B \sum_{j \neq i} \frac{\partial G}{\partial v_i}(v_i^*, v_j^*, \delta_{ij}) f[G(v_i^*, v_j^*, \delta_{ij})] \left\{ \Pi_{k \neq j,i}[1 - F(G(v_i^*, v_k^*, \delta_{ik}))] \right\} = 0. \tag{6.5}$$

To derive some analytical results is quite a difficult task. Therefore, we rely on numerical simulations to provide evidence that the results of Proposition 6.1 hold in the case with more than two competing funds. In our basic simulations, we assume that $m(v) = 1 - (1 - v)^2$ and $N = 3$. In such a case $\hat{v} = 1$ and $\bar{v} = 5/4$. We denote v_w, v_s and v_l the risk levels undertaken in the second period by the funds ranked first, second and third at the end of the first period, respectively. We obtain the following results for 100 observations[3]. For each observation, we have $v_w < v_s < v_l$, $v_w < \hat{v}$ and $v_l > \hat{v}$. This means that *(i)* the risk level chosen in the second period is negatively correlated with the interim performance and *(ii)* the interim winner (w) always decreased his risk level in the second period, while the interim loser (l) always increased his risk level. The fund ranked second either increased or decreased its risk level depending on the performance of the two other funds. Aggregate results from the simulations are given in the following table.

	v_w	v_s	v_l
Average	0.860	1.070	1.390
Std Dev	0.066	0.065	0.668
Max	0.950	1.193	3.501
Min	0.715	0.985	1.045

We observe that, on average, the fund ranked second increased its risk level in the second period.

A final observation is that, on average, the interim loser acts more in the interest of investors that other funds. Its average risk level in the second period (1.39) is the closest to the risk level maximizing expected return (1.25). This confirms the remarks made about the results of Proposition 6.1 that the interim loser may act more in the interest of investors than the interim winner.

[3]For each observation, three independent returns of a random variable normally distributed with mean $m(\bar{v})$ and variance $\bar{v}$ are generated. Then, the system of equations (6.5) is solved numerically under the second-order-condition constraints.

6.6 Empirical evidence

6.6.1 Data

We use mutual fund data from Morningstar's April 1999 Principia Pro Data Disk. The data set contains historical monthly returns, inception date, and various fund characteristics (e.g., fund sizes, expense ratios, minimum investment requirements, etc.). In our analysis, we use a sample of 3096 US diversified equity funds (excluding funds closed to the public), which were active in April 1999. These funds belong to one of five stated objective categories: aggressive growth, growth, growth-and-income, equity-income, and small company. The sample period is from January 1980 to December 1998.

It should be noted that our data set is survivorship biased, excluding funds that disappeared before 1999. As noted by Brown, Harlow, and Starks (1996), if the probability of disappearance is higher for underperforming funds (see, e.g., Brown and Goetzmann, 1993), this bias would be against finding a negative relationship between fund performance and risk, which our model predicts. In addition, our results will be shown to be qualitatively similar in the last years covered in our data set, when the survivorship bias is minimal.

Table 6.1 reports summary statistics of the overall sample as well as objective categories in 1980-1998. During the sample period, an average diversified equity fund realized an annual return of 17.5%, which had standard deviation of about 16% per year. In line with our expectations, aggressive growth and small company funds had the highest total risk, approximately 21% and 19% per year, respectively. However, the highest return was achieved by less risky growth and growth-and-income funds. In all categories, funds, on average, underperformed, according to a one-factor model with the market factor. Jensen's alpha ranges from -0.14% to -0.58% per year for equity-income and small company funds, respectively. Based on the data from the beginning of 1999, an average fund had a six-year performance record and controlled about $610 million in assets.

To illustrate the difference between interim and end-of-the-year rankings, Table 6.2 reports the nine-month return rankings of funds that had top performance over the calendar year. As expected, funds highly ranked after the first three quarters of the year are most likely to top the annual rankings. For example, in the growth category, the top interim performer became the winner of the annual tournament in eleven years

Table 6.1: **Summary statistics of US diversified equity funds**

The table presents summary statistics of the US diversified equity funds calculated for the whole sample and for the aggressive growth, growth, growth-and-income, equity-income, and small company categories. Jensen's alpha, beta, and unsystematic risk are calculated on the basis of the one-factor model with the market factor. Total risk and unsystematic risk are measured as the standard deviation of the returns and the return residuals in the market model, respectively. Mean and standard deviation of fund return, total risk, Jensen's alpha, beta, and unsystematic risk (all annualized) are calculated over the sample period January 1980 - December 1998. Mean and standard deviation of fund size and age are calculated on the basis of cross-sectional data as of beginning of 1999.

		All funds	Ag. Gr.	Growth	Gr. Inc.	Eq. Inc.	Sm. Co.
Total return, %	Mean	17.49	16.64	18.52	18.30	16.53	14.28
	S.D.	15.17	20.00	15.20	12.50	11.72	17.53
Total risk, %	Mean	15.97	20.98	16.47	13.27	11.30	19.06
	S.D.	7.54	8.51	7.26	5.91	4.96	8.37
Jensen's alpha, %	Mean	-0.29	-0.43	-0.27	-0.17	-0.14	-0.58
	S.D.	0.93	1.26	0.84	0.55	0.62	1.41
Beta	Mean	0.95	1.12	0.99	0.90	0.76	0.96
	S.D.	0.27	0.35	0.25	0.18	0.17	0.35
Unsystematic risk, %	Mean	2.08	3.47	2.08	1.12	1.21	3.41
	S.D.	1.41	1.57	1.15	0.71	0.53	1.54
Size, $mln	Mean	609.88	668.47	549.98	1072.19	640.37	190.33
	S.D.	2812.82	2114.02	2030.86	4800.27	2545.32	535.73
Age, years	Mean	6.26	7.85	6.02	7.81	6.56	4.54
	S.D.	9.75	10.26	9.45	12.69	8.76	5.48

out of nineteen. However, sometimes funds ranked as low as 79 out of 191 or 15 out of 19 topped the annual rankings. Thus, the contest for the top annual ranking is not limited to a few funds with best year-to-date performance, and even funds ranked relatively low at the interim stage still have a chance to win the annual tournament.

6.6.2 Tested hypothesis

Assume that manager i has an objective function given in (6.4), i.e., he receives a bonus if fund i's two-period return is the highest among N funds in the category. In this case,

Table 6.2: **Interim performance rankings of top funds**

The table presents interim category return rankings (after the first nine months of the year) of funds with the highest annual returns within a given objective category (aggressive growth, growth, growth-and-income, equity-income, or small company) in each year from 1980 to 1998. The number of funds in a given category and in a given year is in the parentheses.

Year	Ag. Gr.		Growth		Gr. Inc.		Eq. Inc.		Sm. Co.	
1980	1	(16)	1	(98)	1	(59)	2	(15)	1	(13)
1981	2	(18)	1	(104)	1	(59)	1	(15)	2	(14)
1982	1	(21)	1	(109)	2	(62)	2	(15)	1	(16)
1983	1	(21)	8	(115)	1	(68)	1	(15)	1	(20)
1984	1	(24)	1	(125)	1	(72)	1	(17)	2	(28)
1985	1	(30)	1	(138)	7	(82)	15	(19)	2	(33)
1986	5	(30)	2	(157)	1	(93)	1	(21)	1	(41)
1987	1	(35)	10	(178)	2	(105)	5	(25)	3	(48)
1988	1	(39)	1	(203)	1	(120)	4	(35)	1	(65)
1989	3	(39)	7	(217)	1	(132)	5	(43)	1	(73)
1990	12	(39)	1	(235)	2	(140)	5	(45)	1	(77)
1991	1	(43)	1	(256)	2	(159)	1	(51)	4	(84)
1992	1	(43)	2	(285)	79	(191)	1	(56)	1	(106)
1993	3	(49)	1	(377)	1	(235)	2	(71)	1	(142)
1994	2	(65)	2	(502)	3	(309)	1	(93)	7	(206)
1995	1	(76)	7	(649)	1	(395)	18	(126)	2	(283)
1996	5	(97)	1	(815)	1	(480)	1	(142)	8	(362)
1997	1	(127)	4	(1109)	5	(592)	1	(178)	2	(461)
1998	1	(145)	1	(1440)	12	(699)	1	(211)	2	(601)

given the information about the first-period fund performance (denoted Info_1, hereafter), the objective of manager i is to choose the amount of risk in the second period so as to maximize

$$E(C_i|\text{Info}_1) = B\text{Prob}\left(R_{i,1}R_{i,2} > \text{max}_{j\neq i}R_{j,1}R_{j,2}|\text{Info}_1\right)$$

where $R_{i,t}$ is fund i's return in period t. The higher fund i's interim relative performance, the higher the probability of fund i outperforming the other funds at the end of the second period and receiving the bonus. In case of two funds, a fund's interim relative performance can be described by one variable: the difference between its own return and the return of the competing fund. Our theoretical model predicts that the fund's total risk in the second period decreases in this variable (see Proposition 6.1). In case of $N > 2$ funds, fund i's choice of risk in the second period depends on $N-1$ variables:

the differences between fund i's return and the returns of other funds over the first period. In Section 6.5, using simulations, we provided evidence that a general negative relationship between the fund relative performance in the first period and the total risk chosen in the second period holds in case of more than two funds.

For the empirical analysis, the $(N-1)$-dimensional information about the relative performance of a fund over the first period (Info_1) will be summarized by one interim relative performance measure. For the sake of robustness, we use several different specifications of this measure. All of them are non-decreasing functions of the differences between fund i's return and the returns of other funds over the first period, which is taken to be the first three quarters of the year.

Our first measure is fund i's interim category return rank defined as

$$RANK_{i,1} = \frac{1}{N-1}\sum_{j=1}^{N} I_{\{R_{i,1}>R_{j,1}\}},$$

where $I_{\{\}}$ is an indicator function and N is the number of funds in the fund i's category. By construction, $RANK$ ranges from 0 for the worst interim performer to 1 for the top interim performer in the category.

The second measure we use is fund i's interim category-adjusted return:

$$RADJ_{i,1} = R_{i,1} - R_1^{cat},$$

where R_1^{cat} represents the median return in the fund i's category over the first nine months of the year.

Our last variable (denoted $PROB_{i,1}$) measures the probability of fund i finishing the year ranked first in its category (i.e., having the highest annual return in the category), conditional on its interim performance and provided that funds do not change their strategies in the second part of the year and that market conditions do not change. Since we cannot calculate the probability of fund i having the maximum two-period return analytically, we estimate this probability from simulations. The simulation procedure is based on the market model and fund-specific parameters estimated during the first nine months of the year (see Appendix 6.C for a detailed description). Note that, by construction, $PROB$ lies strictly between 0 to 1 and is increasing with fund's interim performance.

In our empirical analysis, we examine whether a fund's choice of risk in the last quarter of the year is negatively related to its interim relative performance measured by the three variables defined above.

6.6.3 Methodology

As discussed in Chapter 5, standard methodology used in the literature (see, e.g., Brown, Harlow, and Starks, 1996, and Koski and Pontiff, 1999) has not produced significant empirical evidence of strategic risk taking by fund managers, which is robust to cross-correlation effects in fund returns. Therefore, in this paper we develop a new, more powerful empirical methodology to examine changes in fund risk.

A manager can influence the level of the fund's total risk in two ways: by changing the fund's factor loadings or the level of the idiosyncratic risk. Testing the model's predictions about the total risk, we should take into account that the fund's total risk may increase or decrease due to the change in market volatility even when its factor betas remain the same. Busse (2001) reports that about 90% of the change in fund standard deviation between the first six months and the last six months of the year arises from changes in the volatility of the common risk (market, size, book-to-market, and momentum) factors and only about 10% from the deliberate actions of fund managers. This will not invalidate tests based on total risk, if all funds have the same factor betas. However, in the case that funds' factor loadings differ from each other, tests based on total risk may produce biased results. There is extensive evidence in the literature (see, e.g., Brown and Goetzmann, 1997) that there are consistent differences between the risk exposures (in particular, market betas) of US diversified equity funds that compose our sample and that these differences are significant not only across, but also within categories. This suggests that changes in fund risk due to fund managers' strategic actions can be better measured by the changes in fund risk exposures.

In this paper, we focus our analysis on the within-year strategic changes in fund exposures to the market factor, which appears to be the most important determinant of fund total risk. Since systematic risk constitutes about 80% of the fund's total risk (see Table 6.1), an increase in the fund's market beta typically results in an increase in the fund's total risk. Due to limitations of our data (fund monthly returns), we do not investigate strategic changes in fund unsystematic risk here. Similarly to tests of the total risk, tests of the unsystematic risk should account for the differences in fund loadings with respect to other factors (e.g., size and momentum) to produce unbiased results. Thus, we investigate whether a fund's systematic risk in the last quarter of the year is related to its relative performance over the first three quarters of the year.

Assume that fund returns over period t ($t = 1$ and $t = 2$ correspond to the first nine months and the last three months of the year, respectively) are generated from a one-factor model with the market factor:

$$R_{i,t} - R_t^f = \alpha_{i,t} + \beta_{i,t}(R_t^m - R_t^f) + \varepsilon_{i,t}, \tag{6.6}$$

where $E(\varepsilon_{i,t}) = 0$ and $E(\varepsilon_{i,t}\varepsilon_{i,s}) = 0$ $(t \neq s)$. $R_{i,t}$, R_t^m, and R_t^f represent the fund i's return, the S&P500 return, and the one-month T-bill rate accumulated over period t, while $\alpha_{i,t}$ and $\beta_{i,t}$ denote the Jensen's alpha and market beta of fund i in period t, respectively.

According to our model, each fund i follows a consistent risk policy with constant beta $\beta_{i,1}$ in the first period (in our setting, the first nine months of the year). In the second period (the last quarter of the year), fund i modifies its beta depending on its interim relative performance $PERF_{i,1}$:

$$\beta_{i,2} = \beta_{i,1} + \gamma PERF_{i,1} + u_{i,2}, \tag{6.7}$$

where $PERF$ is measured as $RANK$, $RADJ$, or $PROB$ over the first three quarters of the year. Substituting (6.7) to (6.6) for $t = 2$ and assuming that fund Jensen's alphas (managerial skills) do not change during the year (i.e., $\frac{1}{3}\alpha_{i,1} = \alpha_{i,2}$), we obtain

$$R_{i,2} - R_2^f - \frac{1}{3}\alpha_{i,1} - \beta_{i,1}(R_2^m - R_2^f) = \gamma PERF_{i,1}(R_2^m - R_2^f) + \varepsilon_{i,2}^*, \tag{6.8}$$

where the residuals $\varepsilon_{i,2}^* = u_{i,2}(R_2^m - R_2^f) + \varepsilon_{i,2}$ are assumed to have zero expectation and be uncorrelated over time. We also assume that $u_{i,2}$ and $\varepsilon_{i,2}$ are uncorrelated, i.e., fund managers do not possess a timing ability.

We estimate the model parameters in two stages. First, we estimate $\alpha_{i,1}$ and $\beta_{i,1}$ on the basis of fund monthly returns during the first nine months of the year. This allows us to compute the market-model residuals over the last quarter of the year (the left-hand side of (6.8)), which would be obtained under the null hypothesis that funds do not change their systematic risk during the year. In the second stage, we estimate γ, using a panel regression approach and the Fama-MacBeth approach.

Our main results are based on the panel regression approach, which is applied to observations pooled over all years in the sample. We run a panel regression with fixed time effects, including year dummies in (6.8). We compute weighted least squares estimates with year-specific weights estimated on the basis of the OLS residuals (i.e., with the variance of the residuals modelled as a function of year dummies). As noted by Busse

(2001), neglecting cross-correlation in fund returns may lead to spurious inference due to the underestimated standard errors. In order to account for the cross-correlation effects, we calculate empirical p-values based on simulations under the null hypothesis of no strategic risk-taking. The simulation procedure is constructed as in Chapter 5. For each month, we simulate the vector of fund returns from a multivariate normal distribution with a mean vector and variance matrix that are estimated from the observed monthly fund returns in a given year. Repeating this process 1000 times, we obtain empirical p-values.[4]

In addition, we estimate the model (6.8) in each of the nineteen annual tournaments from 1980 to 1998. The corresponding p-values are calculated using the same simulation procedure as before, accounting for the cross-correlation in fund returns. Based on the time series of nineteen values of γ corresponding to the annual tournaments, we also calculate the overall Fama-MacBeth estimates of γ.

6.6.4 Results

Table 6.3 presents the results based on model (6.8). Panel A reports the γ coefficients and the corresponding simulated p-values, based on the panel regression approach. The coefficients on all three relative performance measures are negative, being highly significant for $RADJ$ and $PROB$ (p-values are below 1%), and marginally significant for $RANK$ (the p-value is below 10%). The results are significant not only statistically, but also economically. A 40% move (e.g., a move from 20th to 60th return percentile) in category rankings or a 9% change in fund category-adjusted return over the first three quarters of the year are associated with a subsequent change in beta in the last quarter of the year of about 0.1. The same change in beta is also caused by 14% change in $PROB$.

Similar results are obtained using the Fama-MacBeth approach (see Table 6.3, Panel B). The coefficients have approximately the same magnitude as those based on the panel approach and are all highly significant. These results demonstrate that systematic risk chosen by fund managers in the last quarter of the year is negatively related to fund

[4]One can also use a bootstrap approach, as in Busse (2001), to obtain empirical p-values adjusted for the cross-correlation in fund returns. In his sample, the bootstrapped p-values were not materially different from simulated p-values.

Table 6.3: **Relationship between fund interim performance and changes in systematic risk in 1980-1998**

The table documents the relationship between fund performance over the first nine months of the year and subsequent changes in systematic risk. The table reports the γ coefficients and corresponding p-values from the model (6.8) based on the panel regression approach (see Panel A) and the Fama-MacBeth approach (see Panel B). The results are based on the sample of US diversified equity funds in January 1980 - December 1998.

Panel A. Results based on the panel regression approach

Variable	Coef.	p-value
$RANK$	-0.261	0.096
$RADJ$	-0.012	0.015
$PROB$	-0.723	0.000

Panel B. Results based on the Fama-MacBeth approach

Variable	Coef.	p-value
$RANK$	-0.308	0.000
$RADJ$	-0.014	0.000
$PROB$	-0.962	0.000

interim relative performance, which is consistent with Proposition 6.1.

To illustrate the findings in more details, Table 6.4 presents the results of each of the annual tournaments from 1980 to 1998. The pattern of the negative relationship between fund interim performance and subsequent changes in systematic risk is consistent across years. For example, only in three out of nineteen years, the coefficient on $RADJ$ is positive. Most of the significant annual coefficients are also negative. For instance, only one from six coefficients on $PROB$ that are significant at 5% level is positive.

One may expect that strategic behavior is more pronounced among small and young funds, for which it should be easier to change the riskiness of the portfolios. We estimated our model including the interaction terms of fund interim performance and size as well as age in (6.7). However, we did not find any significant differences in systematic risk policy during the last quarter of the year between small and large funds as well as young and old funds.

We also estimated the model (6.8) separately for October, November, and December,

Table 6.4: **Relationship between fund interim performance and changes in systematic risk: annual tournaments in 1980-1998**

The table documents the relationship between fund performance over the first nine months of the year and subsequent changes in systematic risk. The second column reports the total number of funds in the sample in a given year. Columns three to five report the estimated γ coefficients from the model (6.8) for each of the annual tournaments in 1980 - 1998. An asterisk denotes that the corresponding p-value is below 5%.

Year	No. of funds	$RANK$	$RADJ$	$PROB$
1980	201	-0.293	-0.012	-0.589
1981	210	0.041	0.002	0.148
1982	223	-0.068	-0.004	0.092
1983	239	-0.461	-0.023	-1.205*
1984	266	0.221	0.013*	0.428*
1985	302	-0.136	-0.008	-0.765
1986	342	-0.296	-0.017	-1.083*
1987	391	-0.773	-0.028	-1.164
1988	462	-0.349	-0.015	-0.778
1989	504	-0.567*	-0.022*	-0.598
1990	536	-0.030	0.002	0.065
1991	593	-0.447	-0.016	-1.661*
1992	681	-0.312	-0.016	0.255
1993	874	-0.380	-0.018	-1.492*
1994	1175	-0.093	-0.005	-0.351
1995	1529	-0.534	-0.024	-1.140
1996	1896	-0.697	-0.041*	-3.250*
1997	2467	-0.313	-0.012	-1.115
1998	3096	-0.022	-0.001	-0.089

examining whether fund beta in a given month is related to its year-to-date performance. For each of the three months, the relationship between fund interim performance and subsequent change in market beta is negative, being the strongest in statistical terms for October. A statistically weaker results for November and December may be explained by the window-dressing effects documented in, e.g., Carhart et al. (2002).

In addition, we estimated the model (6.8) measuring fund performance relative to an asset class of all diversified equity funds rather than to funds with the same stated objective. Fund managers may be interested in maximizing their asset class relative performance, since it appears to have as strong impact on fund flows as relative performance with respect to the stated objective category (see Chapter 4). We found a similar negative relationship between fund interim performance relative to diversified equity funds and systematic risk in the last quarter of the year. However, this relationship was somewhat weaker in statistical terms than for fund performance relative to the stated objective category.

6.7 Conclusion

The nature of the competition in the money management industry generates relative performance objectives for mutual fund managers. In this chapter, we study how ranking objectives (as in a tournament) influence portfolio decision of a fund manager. In a two-period setting, we show how interim ranking influences the riskiness of the investment strategy chosen by managers in both periods. In the first period, managers choose the same risk level but do not maximize their expected return. In the second period, the interim loser increases risk, while the interim winner decreases risk relative to the first period. Furthermore, the level of risk undertaken by the interim loser is increasing with the difference in interim performances. Using simulations, we demonstrate that the negative relationship between interim performance and risk chosen in the second period also holds in the case of more than two competing funds.

Then, we provide empirical evidence that fund managers' risk-taking behavior is consistent with ranking objectives. Specifically, we find that fund systematic risk in the last quarter of the year is negatively related to the interim category-relative performance, which is consistent with our theoretical predictions. Our statistical tests take into account the presence of the cross-correlation in fund returns, which was not the case in the previous empirical studies of fund risk taking behavior (see, e.g., Brown, Harlow, and Starks, 1996, and Koski and Pontiff, 1999) and could lead to spurious inference (see Busse, 2001).

Finally, our results suggest that investors may be better off taking into account not only performance rankings, but also the difference in performances when selecting between funds. Such allocation rule would "linearize" managers' incentives and mitigate the adverse incentives of fund managers.

Appendix 6.A Proof of Propositions 6.1 and 6.2

Proof of Proposition 6.1: An equilibrium in pure strategies in the period 2 subgame is a pair (v_l^*, v_w^*) such that

$$\tfrac{\partial H_w}{\partial v_w}(v_w^*, v_l^*, \delta) = 0, \qquad \tfrac{\partial H_l}{\partial v_l}(v_l^*, v_w^*, \delta) = 0 \tag{6.9}$$

and

$$\tfrac{\partial^2 H_w}{\partial v_w^2}(v_w^*, v_l^*, \delta) < 0, \qquad \tfrac{\partial^2 H_l}{\partial v_l^2}(v_l^*, v_w^*, \delta) < 0. \tag{6.10}$$

The system of first-order conditions is equivalent to

$$\delta + m(v_w^*) - m(v_l^*) = 2(v_l^* + v_w^*)m'(v_w^*) \tag{6.11}$$

and

$$\delta + m(v_w^*) - m(v_l^*) = -2(v_l^* + v_w^*)m'(v_l^*). \tag{6.12}$$

Conditions (6.11) and (6.12) imply that

$$m'(v_l^*) = -m'(v_w^*). \tag{6.13}$$

Hence, $v_l - \hat{v}$ and $v_w - \hat{v}$ are of opposite signs.

Assume that $m'(v_w^*) < 0$. From (6.11), this implies that $m(v_w^*) < m(v_l^*)$. Given that the interim winner's objective is to maximize $F(G(v_w, v_l, \Delta))$, we deduce that he can increase his probability of winning the contest by choosing $v = v_l^*$. Therefore, there exists a deviation that increases the probability of winning the contest. Hence, there cannot be an equilibrium with $m'(v_w^*) < 0$.

We now show that if $m'(v_w^*) > 0$ then the system of equations (6.11) and (6.12) has a unique solution. Let

$$\mathcal{H}(v_w^*, v_l^*, \delta) = \delta + m(v_w^*) - m(v_l^*) - 2(v_l^* + v_w^*)m'(v_w^*).$$

Equation (6.11) implies that in equilibrium $\mathcal{H}(v_w^*, v_l^*, \delta) \equiv 0$. Now, since $m'(v_w^*) = -m'(v_l^*)$,

$$\frac{d\mathcal{H}}{v_w}(v_w^*, v_l^*, \delta) = -m'(v_w^*)\left(1 - \frac{dv_l^*}{dv_w^*}\right) - 2m''(v_w^*)(v_l^* + v_w^*)$$

with $\frac{dv_l^*}{dv_w^*} = \frac{m''(v_w^*)}{m''(v_l^*)}$. Since $m''(\cdot) < 0$, $m'''(\cdot) \geq 0$ and $v_l^* > v_w^*$, it implies that $\frac{dv_l^*}{dv_w^*} > 1$. Therefore, $\mathcal{H}$ is monotonically increasing in v_w^* with $\lim_{v_w^* \to \hat{v}} \mathcal{H}(v_w^*, v_l^*, \delta) = \delta$ and

$\lim_{v_w^* \to 0} \mathcal{H}(v_w^*, v_l^*, \delta) < 0$. Therefore, the equation $\mathcal{H}(v_w^*, v_l^*, \delta) = 0$ has a unique solution and there exists a unique equilibrium such that $v_w^* < \hat{v} < v_l^*$. The proof that v_w^* and v_l^* are increasing and decreasing in δ, respectively, follows directly from $m'(v_l^*) = -m'(v_w^*)$ and the strict concavity of $m(\cdot)$.

From conditions (6.11) and (6.12), we deduce that

$$d\delta + m'(v_w)dv_w - m'(v_l)dv_l = 2(dv_w + dv_l)m'(v_w) + 2(v_w + v_l)m''(v_w)dv_w, \tag{6.14}$$

$$d\delta + m'(v_w)dv_w - m'(v_l)dv_l = -2(dv_w + dv_l)m'(v_l) - 2(v_w + v_l)m''(v_l)dv_l. \tag{6.15}$$

This implies that

$$m''(v_w)dv_w = m''(v_l)dv_l. \tag{6.16}$$

In turn, this implies that dv_l and dv_w are of opposite signs. Furthermore, from (6.13), (6.14) and (6.16), we obtain that

$$d\delta = dv_w \left[2(v_l + v_w)m''(v_w) + m'(v_w) \left(1 - \frac{m''(v_w)}{m''(v_l)} \right) \right]. \tag{6.17}$$

Given the assumption that $m'''(\cdot) > 0$ and the result that $v_l^* > v_w^*$ in equilibrium, it follows that $m''(v_w^*)/m''(v_l^*) < 1$. Hence, v_w^* and v_l^* are decreasing and increasing in δ, respectively.. $\square$

Proof of Proposition 6.2: Let $\delta_{ij} = r_{i,1} - r_{j,1}$ $(i, j = 1, 2,\ i \neq j)$. From the proof of Proposition 6.1, we know that a manager who is leading after the first period has a probability strictly larger than 1/2 of winning the contest. Now, if manager 1 chooses $v_1 = \hat{v}$ in the first period, then for any $v_2 \neq \hat{v}$ chosen by manager 2, $\text{Prob}(\delta_{2,1} > 0) < 1/2$, while if manager 2 chooses $v_2 \neq \hat{v}$ in the first period, then $\text{Prob}(\delta_{2,1} > 0) = 1/2$. Hence, $\hat{v}$ is a best reply to $\hat{v}$.. $\square$

Appendix 6.B Correlated returns

In this appendix, we analyze the case in which managers choose among portfolios with correlated returns. To do so, we modify the model of Section 6.3 in the following way. Assume that a safe asset (S) with return normalized to 1, and two risky portfolios are available. These two portfolios (hereafter, p_a and p_b) have returns (R_a and R_b) independently and normally distributed with variances $v_a = \hat{v}$ and $v_b > \hat{v}$ and means $m_a = m(v_a)$ and $m_b = m(v_b)$ (with $m_a > 1$ and $m_b > 1$), respectively; the function $m(\cdot)$ and $\hat{v}$ being as defined in Section 6.3. Therefore, $m_b < m_a$.

Denote l the interim loser and w the interim winner. In the second period, manager j ($j = l, w$) chooses an allocation $(\theta_{aj}, \theta_{bj})$, θ_{aj} and θ_{bj} being invested in portfolio p_a and p_b, respectively, and $(1 - \theta_{aj} - \theta_{bj})$ being invested in asset S. It follows that the return of manager j in the second period is

$$R_{j,2} = 1 + \theta_{aj}(R_a - 1) + \theta_{bj}(R_b - 1).$$

For tractability, we restrict the set of choices to $\theta_{aj} \geq 0$, $\theta_{bj} \geq 0$ and $\theta_{aj} + \theta_{bj} < 1$. This implies that shortselling the safe asset or the two risky portfolios is forbidden.

The main difference with Section 6.3 is that now returns are correlated, and their covariance is endogenous:

$$\text{cov}(R_{l,2}, R_{w,2}) = \theta_{al}\theta_{aw}v_a + \theta_{bl}\theta_{bw}v_b.$$

Let $R_{1w}/R_{1,l} = \Delta$.[5] It is straightforward that the best reply of the interim winner is to choose the same allocation as the loser since in such a case, he wins the contest with probability 1. Conversely, the objective of the interim loser is to choose an allocation that generates a return correlated as little as possible with the return of the interim winner. It follows that such a game has only equilibria in mixed strategies. For some of these equilibria, we can derive results about the relative amount of risk undertaken by the two managers.

Proposition 6.3 *Assume that $A = 0$ and consider any equilibrium such that* (i) *managers only invest in the risky portfolios (i.e., $\theta_{aj} + \theta_{bj} = 1$, $j = l, w$) and* (ii) *man-*

[5] In the previous sections, it was assumed that managers are identical ex-ante. Here, there always exists an equilibrium such that $\Delta = 1$ with probability 1 given the two risky portfolios available. Therefore, $\Delta > 1$ requires that managers did not choose the same portfolio in period 1. One possibility is that they were heterogeneously informed in period 1, while this is not the case in period 2.

agers randomize between the two same allocations $(\theta_{aj}, \theta_{bj}) = (\theta', 1-\theta')$ *or* $(\theta_{aj}, \theta_{bj}) = (\theta'', 1-\theta'')$ $(j = l, w)$ *with* $\theta' > \theta''$. *Denote* q_j *the equilibrium probability that manager* j *chooses* $\theta_{aj} = \theta'$. *Then, in such an equilibrium, the interim loser takes, on average, more risk than the interim winner: the interim loser chooses* θ' *with a lower probability than the interim winner, i.e.,* $q_l < q_w$.

Proof: Consider any equilibrium that satisfies conditions *(i)* and *(ii)*. Given the equilibrium strategy of the interim loser (i.e., the probability q_l with which he chooses θ'), the interim winner is indifferent between the two pure strategies. This implies that

$$\begin{aligned} & q_l \text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta', \theta_{al} = \theta'\right) + (1-q_l)\text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta', \theta_{al} = \theta''\right) = \\ & q_l \text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta'', \theta_{al} = \theta'\right) + (1-q_l)\text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta'', \theta_{al} = \theta''\right). \end{aligned} \tag{6.18}$$

Given that

$$\text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta', \theta_{al} = \theta'\right) = \text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta'', \theta_{al} = \theta''\right) = 0,$$

it follows that

$$q_l = \frac{\text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta', \theta_{al} = \theta''\right)}{\text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta', \theta_{al} = \theta''\right) + \text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta'', \theta_{al} = \theta'\right)}.$$

Proceeding similarly, we find that

$$q_w = \frac{\text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta'', \theta_{al} = \theta'\right)}{\text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta', \theta_{al} = \theta''\right) + \text{Prob}\left(R_{l,2}/R_{w,2} > \Delta | \theta_{aw} = \theta'', \theta_{al} = \theta'\right)}.$$

Let $r_{2,i} = \log(R_{2,i})$ $(i = l, w)$ and $\delta = \log(\Delta)$. Then, $q_w > q_l$ is equivalent to

$$\text{Prob}\left(r_{l,2} - r_{w,2} > \delta | \theta_{aw} = \theta'', \theta_{al} = \theta'\right) > \text{Prob}\left(r_{l,2} - R_{w,2} > \delta | \theta_{aw} = \theta', \theta_{al} = \theta''\right).$$

This is equivalent to

$$\Phi\left(\frac{\delta - (\theta' - \theta'')(m_a - m_b)}{\sqrt{(\theta' - \theta'')^2 (v_a + v_b)}}\right) < \Phi\left(\frac{\delta - (\theta'' - \theta')(m_a - m_b)}{\sqrt{(\theta' - \theta'')^2 (v_a + v_b)}}\right),$$

where Φ is cdf of the standard normal distribution. Given that $\theta' > \theta''$, this last inequality always holds. $\square$

If managers do not buy the risk-free bond, then the larger θ_{bj}, the larger the amount of risk taken by manager j. Proposition 6.3 states that in any equilibrium such that managers do not buy the risk-free bond and choose among the same two allocations, the interim loser takes more risk than the interim winner, on average.

This result is different from Taylor (2000) for one main reason. Taylor considers an economy with a risk-free asset and *one* risky asset. It follows that an interim winner increasing risk also increases the expected return of his portfolio. He does not face a trade-off between increasing the variance and decreasing the expected return. Conversely, we consider a situation such that managers have the possibility to choose portfolios with low expected return and high variance .

We can derive further results on the interim loser's risk taking incentives.

Proposition 6.4 *Assume that $A = 0$ and the interim winner chooses a portfolio such that $\theta_{aw} + \theta_{bw} = 1$ (i.e., does not buy the risk-free bond). If $\theta_{aw} \leq 1/2$, then the best reply of the interim loser is $\theta_{al} = 1$. If $\theta_{aw} > 1/2$, then the best reply of the interim loser is $\theta_{bl} = 1$.*

Proof: Proceeding as in the previous section, one shows that the objective of interim loser is to maximize

$$H(\theta_{al}, \theta_{bl}, \theta_{aw}, \theta_{bw}, \delta) = \frac{-\delta + (\theta_{al} - \theta_{aw})(m_a - 1) + (\theta_{bl} - \theta_{bw})(m_b - 1)}{\sqrt{(\theta_{aw} - \theta_{al})^2 v_a + (\theta_{bw} - \theta_{al})^2 v_b}}$$

with respect to θ_{al} and θ_{bl} under the constraint that $\theta_{al} + \theta_{bl} \leq 1$. First, we show that there cannot be an interior solution to this problem. To see this, assume that the interim loser chooses $\theta_{bl} \in (0, 1)$.

$$\frac{\partial H}{\partial \theta_{al}} = \frac{(m_a - 1)(\theta_{bw} - \theta_{bl})^2 v_b - v_a(\theta_{al} - \theta_{aw})\left[(\theta_{bl} - \theta_{bw})(m_b - 1) - \delta\right]}{\left((\theta_{aw} - \theta_{al})^2 v_a + (\theta_{bw} - \theta_{al})^2 v_b\right)^{3/2}}.$$

Therefore, if $\theta_{bl} < \theta_{bw}$, for any θ_{al}, $\partial H/\partial \theta_{al} > 0$. It implies that the interim loser chooses $\theta_{al} = 1 - \theta_{bl}$. Now if $\theta_{bl} > \theta_{bw}$, then it implies that $\theta_{al} < \theta_{aw}$ (since, by assumption, $\theta_{aw} + \theta_{bw} = 1$).

$$\frac{\partial H}{\partial \alpha_{bl}} = \frac{(m_b - 1)(\theta_{aw} - \theta_{al})^2 v_a - v_b(\theta_{bl} - \theta_{bw})\left[(\theta_{al} - \theta_{aw})(m_a - 1) - \delta\right]}{\left((\theta_{aw} - \theta_{al})^2 v_a + (\theta_{bw} - \theta_{al})^2 v_b\right)^{3/2}}.$$

If $\theta_{al} < \theta_{aw}$, then for any α_{al}, $\partial H/\partial \theta_{bl} > 0$. It implies that the interim loser chooses $\theta_{bl} = 1 - \theta_{al}$. Therefore, we always have $\theta_{aw} + \theta_{al} = 1$.

This implies that the problem of the interim loser is to choose θ_{al} to maximize

$$K(\theta_{al}, \theta_{aw}, \delta) = \frac{-\delta + (\theta_{al} - \theta_{aw})(m_a - m_b)}{\sqrt{(\theta_{aw} - \theta_{al})^2 (v_a + v_b)}}$$

under the constraint that $\theta_{al} \in [0, 1]$. It is straightforward that this is equivalent to maximizing $|\theta_{al} - \theta_{aw}|$. Therefore, if $\theta_{aw} < 1/2$, the interim loser chooses $\theta_{al} = 1$, while if $\theta_{aw} > 1/2$, the interim loser chooses $\theta_{bl} = 1$. □

This proposition tells us that the best reply of the interim loser to an allocation of only risky portfolios by the interim winner is to choose the allocation of only risky portfolios that minimizes the correlation with the return of the interim winner.

From Propositions 6.3 and 6.4, we derive the following result.

Proposition 6.5 *Assume that $A = 0$. There exists an equilibrium such that*
(i) *the interim winner chooses $\theta_{aw} = 1$ with probability q_w and $\theta_{bw} = 1$ with probability $(1 - q_w)$*
(ii) *the interim loser chooses $\theta_{al} = 1$ with probability q_l and $\theta_{bl} = 1$ with probability $(1 - q_l)$*
(iii) *$q_w > q_l$.*

Proof: As already mentioned, the best reply of the interim winner is to play the same strategy as the interim loser. Furthermore, from Proposition 6.4, we know that $(\theta_{al}, \theta_{bl}) = (1, 0)$ is a best reply to $(\theta_{aw}, \theta_{bw})$ with $\theta_{aw} > 1/2$ and $\theta_{aw} + \theta_{bw} = 1$; and that $(\theta_{al}, \theta_{bl}) = (0, 1)$ is a best reply to $(\theta_{aw}, \theta_{bw})$ with $\theta_{aw} < 1/2$ and $\theta_{aw} + \theta_{bw} = 1$. This implies that there exists an equilibrium in which manager j chooses $(\theta_{aj}, \theta_{bj}) = (1, 0)$ with probability q_j and $(\theta_{aj}, \theta_{bj}) = (0, 1)$ with probability $1 - q_j$ $(j = w, l)$. Proposition 6.3 implies that $q_w > q_l$. □

This proposition states that there exist equilibria such that Proposition 6.3 holds: when both the variance and the covariance of the portfolios are strategic variables, then, on average, the interim loser takes more risk than the interim winner. Hence, the results derived in Section 6.4 still hold (qualitatively) when returns are correlated and their covariance level is a strategic variable.

Appendix 6.C Simulation procedure for the third relative performance measure

The third interim relative performance measure used in this chapter, $PROB_{i,1}$, is the estimate of the probability that fund i has the highest annual return in its category, conditional on fund performance over the first three quarters of the year and given that funds do not change their strategies and that market conditions do not change in the last quarter of the year. This appendix describes the simulation procedure used to compute this measure.

We use a market model (6.6) as a basis for our simulations. We simulate fund last-quarter returns using the distribution parameters estimated on the basis of fund monthly returns during the first three quarters of the year. Specifically, we estimate fund Jensen's alphas and market betas, the mean and variance of the excess market return, and the variance matrix of the market-model residuals (in order to preserve the cross-correlation structure of fund returns). The vector of simulated fund returns over the last quarter of the year is then calculated as a function of fund Jensen's alphas and betas as well as randomly generated values of the excess market return and the market-model residuals.

Formally, for each category consisting of N funds, we simulate the $N \times 1$ vector of fund returns over the last quarter of the year using the following formula:

$$R_2 = \frac{1}{3}(\alpha_1 + \beta_1 RMRF_1 + e_1), \tag{6.19}$$

where α_1 and β_1 denote the $N \times 1$ vectors of Jensen's alphas and market betas of funds estimated over the first three quarters of the year, respectively. The excess market return $RMRF_1$ is generated from a normal distribution with mean and variance calculated on the basis of monthly excess market returns in the first nine months of the year. The vector of residuals e_1 is generated from a normal distribution with zero mean and variance matrix estimated on the basis of monthly market-model residuals in the first three quarters of the year. Note that the excess market returns, Jensen's alphas and market-model residuals are calculated on the three-quarter basis and should be divided by three to obtain a quarterly return in (6.19). The simulated probability of becoming a top fund in the category is based on 1000 replications of this procedure.

Chapter 7

Summary

This thesis consists of two parts, in which we investigate the allocation rules of mutual fund investors and strategies of mutual fund managers. These two aspects are closely related to each other. On the one hand, investors try to select funds that follow an optimal investment policy from their point of view. On the other hand, fund managers are interested in maximizing net fund inflows, when their compensation is linked to the fund's size (see, e.g., Khorana, 1996). The aim of this thesis is to contribute to the understanding of the behavior of mutual fund investors and managers as well as the link between the two.

Chapter 2 provides an introduction to the main topics investigated in the literature on mutual funds. Numerous studies are devoted to the evaluation of mutual fund performance, trying to answer the question whether active fund management adds value. It has been demonstrated that mutual funds, as a group, have negative or neutral performance adjusted for risk and expenses (see, e.g., Gruber, 1996, and Ferson and Schadt, 1996). Several studies show that it is possible to identify funds with consistent superior performance as well as funds with consistent inferior performance, based on such factors as fund past performance and expenses (see, e.g., Kosowski et al., 2000, and Baks, Metrick, and Wachter, 2001). Clearly, fund investors are better off choosing the former and avoiding the latter funds. Indeed, studies of mutual fund flows find a clear positive relationship between past fund performance and their subsequent flows (see, e.g., Sirri and Tufano, 1998). However, this relationship appears to be convex, i.e., flows to top performers are more sensitive to performance than flows to poor performers (see, e.g., Chevalier and Ellison, 1997). Since managerial compensation is typically based on a proportion of the fund's assets (see, e.g., Khorana, 1996), this leads to the convexity in the manager's expected payoff as a function of fund performance. Several studies using game-theoretic framework demonstrate that compensation contracts convex or linear in the fund's benchmark-adjusted performance may induce adverse incentives to fund managers with respect to the choice of effort and risk (see, e.g., Hvide, 1999, and Carpenter, 2000). The existing empirical evidence suggests that fund choice of risk may be related to its past performance (see, e.g., Brown, Harlow, and Starks, 1996). However, most of these results should be taken with caution, since they are based on statistical tests that did not take the auto-correlation and cross-correlation in fund returns into account (see Busse, 2001). Several studies find the evidence of the gaming behavior by fund managers around the year-ends (see, e.g., Musto, 1999, and Carhart, et al., 2002).

In Chapters 3 and 4, we analyze the determinants of mutual fund flows, concentrating on the impact of past performance on fund flows. Chapter 3 provides empirical evidence on the dynamics of the impact of past performance on flows to US growth funds in 1991-1999. We identify significant nonlinearities in the lag structure of the flow-performance relationship. In particular, we observe that performance during the most recent quarter is less important than performance during the remaining three quarters of the first year, suggesting that some investors react to fund performance with a certain lag. The first three years of past performance history account for about 90 percent of the total impact of past performance on flows. In addition, we demonstrate that the return on systematic risk factors has a small additional impact on fund flows, on top of the impact of risk-adjusted returns. This implies that some mutual fund investors are style timers, choosing funds on the basis of raw rather than risk-adjusted performance.

Chapter 4 presents empirical evidence on the impact of different classification systems on flows to US mutual funds in 1993-1999. Specifically, we examine the relationship between flows to US mutual funds and their performance rankings within three types of categories: funds with the same stated objective, funds with the same Morningstar style, and funds within the same asset class. We find that Morningstar style and asset class rankings have a strong positive impact on fund flows on top of the impact of the stated objective rankings. In fact, the asset class ranking appears to be the most important relative performance measure for private investors of domestic stock funds as well as investors of taxable bond and municipal bond funds. Institutional investors of domestic stock funds attach approximately equal weights to the Morningstar style and stated objective rankings. Only in the international stock class, flow-performance sensitivity is the highest for the stated objective ranking. In a joint model of ordinal and cardinal performance measures, the impact of total past returns on fund flows never exceeds the combined impact of performance rankings. Performing a category-specific analysis of the star spillover effect (see, e.g., Nanda, Wang, and Zheng, 2000), we find that the presence of a star fund is typically beneficial for flows to the other funds in its family. Only in case of taxable bond funds, top performance of a star fund "cannibalizes" flows to funds with the same stated objective as the star fund.

Chapters 5 and 6 present the game-theoretic as well as empirical analysis of the behavior of mutual fund managers. In Chapter 5, we consider the statistical tests of the risk taking by mutual fund managers performed in the literature. Busse (2001) notes that autocorrelation in daily fund returns biases volatility estimates, while cross-dependencies in mutual fund returns invalidate the independence assumption underlying the standard statistical tests of fund risk taking. Busse (2001) argues that the evidence in favor of the tournament hypothesis that the within-year changes in risk are related to fund interim performance disappears when auto-correlation and cross-correlation in fund returns are taken into account. In Chapter 5, we contribute to this debate by considering the impact of both auto-correlation and cross-correlation on the tournament tests from an analytical point of view. First, we give analytical expressions for the biases arising in volatility estimates (based on both daily and monthly data) due to first-order autocorrelation effects in the daily fund returns. We show that tests of the tournament hypothesis based on monthly data are in fact more robust to autocorrelation effects than tests based on daily data. Second, to address the impact of cross-correlated fund returns on the tests, we provide explicit conditions under which the tests used in the literature have appropriate size properties. This shows that a specific form of cross-sectional dependence in the mutual fund returns is allowed without affecting the statistical properties of the tests. The crucial condition is that idiosyncratic fund returns should be independent.

In Chapter 6, we study risk taking incentives of mutual fund managers who have ranking objectives (as in a tournament). First, in a two-period model, we analyze the game played by two risk-neutral fund managers with ranking objectives. We show that in the first period, managers choose the same risk level but do not maximize their expected return. In the second period, the interim loser *(i)* increases risk with respect to the first period, while the interim winner decreases risk, *(ii)* the difference in risk undertaken is increasing with the difference in interim performances. Using simulations, we demonstrate that manager's choice of risk in the second period is negatively related to his relative performance over the first period also in the case with more than two competing funds. Second, we empirically test some predictions of the model in a sample of US diversified equity funds in 1980-1998, using a more powerful methodology than in previous studies and accounting for cross-correlation in fund returns. We find evidence that fund choice of systematic risk in the last quarter of the year is negatively related to the performance over the first three quarters of the year, which is consistent with the model.

The analysis of the mutual fund industry conducted in this thesis can be extended in several directions. One interesting topic for further research is the impact of past returns on various systematic risk factors, such as size or momentum, on the behavior and strategies of mutual fund investors and managers. On the one hand, investors may select funds not only on the basis of their past risk-adjusted performance, but also the risk characteristics of their portfolios. For example, they may reward funds with high exposure to the "hot" factors which recently realized high returns. Although Chapter 3 provides preliminary evidence that some investors take fund raw performance into account, further analysis is required to identify separate impact of different systematic risk factors on mutual fund flows. On the other hand, fund managers may pursue a similar "style-timing" strategy of increasing the exposure to the well-performing risk factors. This strategy may help them to minimize the gap in performance with respect to funds that concentrate their investments in "hot" styles and have higher raw returns (see, e.g., Barberis and Schleifer, 2000). Such behavior may be more pronounced for managers of small and more volatile funds, for which it is easier to make significant changes in investment policy, as well as managers of underperforming funds, who are likely to change the fund's strategy in order to improve fund performance and decrease the probability of being fired. Another topic for further investigation is the identification of the calendar-year effects in the dynamic structure of the flow-performance relationship. The question is whether year-to-date performance has a separate impact on fund flows, on top of the impact of past performance measured over a fixed, e.g., one-year, rolling horizon.

References

Admati, Anat R., and Paul Pfleiderer, 1996, Does it all add up? Benchmarks and the compensation of active portfolio managers, *Journal of Business* 70, 323-350.

Baks, Klaas, Andrew Metrick, and Jessica Wachter, 2001, Should investors avoid all actively managed mutual funds? A study in Bayesian performance evaluation, *Journal of Finance* 56, 45-85.

Barber, Brad M., Terrance Odean, and Lu Zheng, 2001, Out of sight, out of mind. The effect of expenses on mutual fund flows, working paper, Graduate School of Management, UC Davis and School of Business Administration, University of Michigan.

Barberis, Nicolas, and Andrei Shleifer, 2000, Style investing, working paper, University of Chicago and Harvard University.

Barra Strategic Consulting Group, 2001, Style is dead. Long live style! August, research report.

Bergstresser, Daniel, and James Potterba, 2002, Do after-tax returns affect mutual fund inflows? *Journal of Financial Economics* 63, 381-414.

Bogle, John C., 1998, The implications of style analysis for mutual fund performance, *Journal of Portfolio Management* 24 (4), 34-42.

Boudoukh, Jacob, Matthew P. Richardson, and Robert F. Whitelaw, 1994, A tale of three schools: Insights on autocorrelations of short-horizon stock returns, *Review of Financial Studies* 7, 539-573.

Brown, Stephen J., and William N. Goetzmann, 1995, Performance persistence, *Journal of Finance* 50, 679-698.

Brown, Stephen J., and William N. Goetzmann, 1997, Mutual fund styles, *Journal of Financial Economics* 43, 373-399.

Brown, Keith C., W. V. Harlow, and Laura T. Starks, 1996, Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry, *Journal of Finance* 51, 85-110.

Busse, Jeffrey A., 2001, Another look at mutual fund tournaments, *Journal of Financial and Quantitative Analysis* 36, 53–73.

Cabral, Luis M.B., 1997, Football, sailing and R&D: Dynamic competition with strategic choice of variance and co-variance, working paper, London Business School.

Capon, Noel, Gavan J. Fitzsimons, and Russ Alan Prince, 1996, An individual level analysis of mutual fund investment decision, *Journal of Financial Services Research* 10, 59-82.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Carhart, Mark M., Ron Kaniel, David K. Musto, and Adam V. Reed, 2002, Leaning for the tape: Evidence of gaming behavior in equity mutual funds, *Journal of Finance* 57, 661-693.

Carpenter, Jennifer N., 2000, Does option compensation increase managerial risk appetite? *Journal of Finance* 55, 2311-2331.

Chamberlain, Gary, and Michael Rothchild, 1983, Arbitrage, factor structure, and mean-variance analysis on large asset markets, *Econometrica* 51, 1281–1304.

Chan, Louis K.C., Hsiu-Lang Chen, and Josef Lakonishok, 2002, On mutual fund investment styles, *Review of Financial Studies* forthcoming.

Chen, Hsiu-lang, and George C. Pennacchi, 1999, Does prior performance affect a mutual fund's choice of risk? Theory and further empirical evidence, working paper, University of Illinois at Chicago and University of Illinois at Urbana-Champaign.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167-1200.

Chevalier, Judith, and Glenn Ellison, 1999a, Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance, *Journal of Finance* 54, 875-899.

Chevalier, Judith, and Glenn Ellison, 1999b, Career concerns of mutual fund managers, *Quarterly Journal of Economics* 114, 389-432.

Christoffersen, Susan E. K., 2001, Why do money fund managers voluntarily waive their fees? *Journal of Finance* 56, 1117-1140.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035-1058.

Das, Sanjiv Ranjan, and Rangarajan K. Sundaram, 2002, Fee speech: Signaling, risk sharing, and the impact of fee structures on investor welfare, *Review of Financial Studies* forthcoming.

Davis, James L., 2001, Mutual fund performance and manager style, *Financial Analysts Journal* 57 (1), 19-27.

Del Guercio, Diane, and Paula A. Tkac, 2001, Star power: The effect of Morningstar ratings on mutual fund flows, working paper 2001-15, Federal Reserve Bank of Atlanta.

Del Guercio, Diane, and Paula A. Tkac, 2002, The determinants of the flow of funds of managed portfolios: Mutual funds versus pension funds, *Journal of Financial and Quantitative Analysis* forthcoming.

Deli, Daniel N., 2002, Mutual fund advisory contracts: An empirical investigation, *Journal of Finance* 57, 109-133.

DiBartolomeo, Dan, and Erik Witkowski, 1997, Mutual fund misclassification, *Financial Analysts Journal* 53 (5), 32-43.

Edelen, Roger M., 1999, Investor flows and the assessed performance of open-end mutual funds, *Journal of Financial Economics* 53, 439-466.

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake, 2002, Incentive fees and mutual funds, *Journal of Finance* forthcoming.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on bonds and stocks, *Journal of Financial Economics* 33, 3-53.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425-461.

Geweke, John, 1978, Temporal aggregation in the multiple regression model, *Econometrica* 46, 643-661.

Godfrey, L.G., 1991, *Misspecification tests in econometrics*, Econometric Society Monographs.

Goetzmann, William N. and Nadav Peles, 1997, Cognitive dissonance and mutual fund investors, *Journal of Financial Research* 20, 145-158.

Gould, C., 1998, Tiny fund sees big opportunities as Euro approaches, *International Herald Tribune*, Dec. 16, 18.

Grinblatt, Mark, and Sheridan Titman, 1989, Adverse risk incentives and the design of performance-based contracts, *Management Science* 35, 807-822.

Gruber, Martin J., 1996, Another puzzle: The growth in actively managed mutual funds, *Journal of Finance* 51, 783-810.

Heinkel, Robert, and Neal M. Stoughton, 1994, The dynamics of portfolio management contracts, *Review of Financial Studies* 7, 351-387.

Hendricks, Darryl, Jay Patel and Richard Zeckhauser, 1993, Hot hands in mutual funds: short run persistence of relative performance, 1974-1988, *Journal of Finance*, 48, 93-130.

Huddart, Steven, 1999, Reputation and performance fee effects on portfolio choice by investment advisers, *Journal of Financial Markets* 2, 227-271.

Hvide, Hans K., 1999, Tournament rewards and risk taking, working paper, Norwegian School of Economics.

Investment Company Institute, 2002, *Mutual fund fact book*.

Ippolito, Richard A., 1992, Consumer reaction to measures of poor quality: Evidence from the mutual fund industry, *Journal of Law and Economics* 35, 45-70.

Ivkovic, Zoran, 2000, Is blood thicker than water: Spillovers in mutual fund families, working paper, Yale School of Management.

Jain, Prem C. and Joanna S. Wu, 2000, Truth in mutual fund advertising: Evidence on future performance and fund flows, *Journal of Finance* 55, 937-958.

Jensen, Michael C., 1969, Risk, the pricing of capital assets, and the evaluation of investment portfolios, *Journal of Business* 42, 167-247.

Khorana, Ajay, 1996, Top management turnover. An empirical investigation of mutual fund managers, *Journal of Financial Economics* 40, 403-427.

Khorana, Ajay, and Henri Servaes, 2001, An examination of competition and investor behavior in the mutual fund industry, working paper, Georgia Institute of Technology and London Business School.

Kim, Moon, Ravi Shukla, and Michael Tomas, 2000, Mutual fund objective misclassification, *Journal of Economics and Business* 52, 309-323.

Kosowski, Robert, Alan Timmerman, Hal White, and Russ Wermers, 2000, Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, working paper, London School of Economics, University of California, San Diego, and University of Maryland.

Koski, Jennifer L., and Jeffrey Pontiff, 1999, How are derivatives used? Evidence from the mutual fund industry, *Journal of Finance* 54, 791-816.

Lettau, Martin, 1997, Explaining the facts with adaptative agents: The case of mutual fund flows, *Journal of Econonic Dynamics and Control* 21, 1117-1147.

Lynch, Antony W., and David K. Musto, 2000, How investors interpret past returns, working paper, New York University and University of Pennsylvania.

Massa, Massimo, 1997, Do investors react to mutual fund performance? An imperfect competition approach, working paper, Yale University.

Meyer, Margaret A., and John Vickers, 1997, Performance comparison and dynamic incentives, *Journal of Political Economy* 105, 547-581.

Musto, David, 1999, Investment decisions depend on portfolio disclosure, *Journal of Finance* 54, 935-952.

Nanda, Vikram, Zhi Wang, and Lu Zheng, 2000, Family values and the star phenomenon, working paper, University of Michigan Business School.

Palomino, Frederic, 2002, Relative performance objectives in financial market, working paper, Tilburg University.

Palomino, Frederic, and Andrea Prat, 2002, Risk taking and optimal contracts for money managers, *RAND Journal of Economics* forthcoming.

Palomino, Frederic, and Harald Uhlig, 2002, Should smart investors buy funds with high returns in the past? working paper, Tilburg University and Humboldt University.

Pastor, Lubos, and Robert F. Stambaugh, 2002, Investing in equity mutual funds, *Journal of Financial Economics* 63, 351-380.

Pozen, Robert C., 1998, The mutual fund business, The MIT Press.

Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589-1622.

Taylor, Jonathan D., 2000, A role for the theory of tournaments in studies of mutual fund behavior, working paper, MIT.

Teo, Melvyn, and Sung-Jun Woo, 2001, Persistence in style-adjusted mutual fund returns, working paper, Harvard University.

Wermers, Russ, 2000, Mutual fund performance: an empirical decomposition into stock picking talent, style transaction costs and expenses, *Journal of Finance* 55, 1655-1695.

Zeckhauser, Richard, Jayendu Patel, and Darryll Hendricks, 1991, Nonrational actors and financial market behavior, *Theory and Decision* 31, 257-287.

Zheng, Lu, 1999, Is money smart? A study of mutual fund investors' selection ability, *Journal of Finance* 54, 901-933.

Samenvatting (Summary in Dutch)

Dit proefschrift bestaat uit twee delen, waarin achtereenvolgens de keuze van beleggers voor een bepaald beleggingsfonds en de strategieën van de beheerders van beleggingsfondsen onderzocht worden. Deze twee aspecten zijn nauw met elkaar verbonden. Enerzijds proberen beleggers beleggingsfondsen te selecteren die een voor hen optimaal investeringsbeleid implementeren. Anderzijds zijn fondsmanagers geïnteresseerd in het maximaliseren van netto fondstoeloop (gedefinieerd als de nieuwe inleg minus de opgevraagde gelden), aangezien hun compensatie gerelateerd is aan de fondsgrootte. Het doel van dit proefschrift is bij te dragen aan een beter begrip van het gedrag van beleggers in een beleggingsfonds, beheerders van een beleggingsfonds en de link tussen beide.

Hoofdstuk 2 geeft een inleiding op de belangrijkste bevindingen uit de literatuur betreffende beleggingsfondsen. Talrijke studies zijn gewijd aan de evaluatie van de prestatie van beleggingsfondsen, in een poging de vraag te beantwoorden of actief fondsbeheer zinvol is. Vrijwel steeds wordt geconcludeerd dat beleggingsfondsen, als groep, slechter of hoogstens even goed als de markt presteren in termen van voor risico en kosten gecorrigeerde rendementen (zie bijvoorbeeld Gruber, 1996, en Ferson & Schadt, 1996). Verscheidene studies tonen evenwel aan dat het mogelijk is zowel fondsen met voortdurend superieure rendementsprestaties als fondsen met voortdurend inferieure prestaties te identificeren. Deze identificatie is dan gebaseerd op factoren zoals eerdere prestaties van het fonds en kosten (zie bijvoorbeeld Kosowski e.a., 2000, en Baks, Metrick en Wachter, 2001). Het mag duidelijk zijn dat fondsbeleggers beter af zijn wanneer zij voor het eerste soort fonds kiezen en de tweede negeren. Inderdaad laten studies van toeloop naar een beleggingsfonds een duidelijk positief verband zien tussen vroegere prestaties van het fonds en de toeloop die daarop volgt (zie bijvoorbeeld Sirri en Tufano, 1998). Deze relatie blijkt echter convex te zijn, d.w.z. stromen naar toppresteerders zijn gevoeliger voor prestaties dan stromen naar slechte presteerders (zie bijvoorbeeld Chevalier en Ellison, 1997). Aangezien het salaris van managers in het algemeen gebaseerd is op een deel van de fondsgrootte (zie bijvoorbeeld Khorana, 1996), leidt dit tot convexiteit van het verwachte salaris van de manager als functie van de fondsprestatie.

Verscheidende studies tonen, gebruikmakend van een speltheoretisch raamwerk, aan dat salariscontracten die convex of lineair gespecificeerd zijn in de normgecorrigeerde fondsprestatie verkeerde incentives voor fondsbeheerders kunnen veroorzaken met betrekking tot hun keuze van inspanning en risico (zie bijvoorbeeld Hvide, 1999, en Carpenter, 2000). Het bestaande empirische bewijs suggereert dat de risico-keuze van een fondsbeheerder gerelateerd kan zijn aan eerdere prestaties van dat fonds (zie bijvoorbeeld Brown, Harlow, en Starks, 1996). Deze resultaten moeten met voorzichtigheid worden beoordeeld, aangezien ze gebaseerd zijn op statistische toetsen die geen rekening hebben gehouden met de autocorrelatie en kruiscorrelatie in fondsrendementen (zie Busse, 2001). Overigens vinden verscheidene studies ook bewijs voor strategisch gedrag van fondsbeheerders rond het einde van het jaar (zie bijvoorbeeld Musto, 1999, en Carhart e.a., 2002).

In Hoofdstuk 3 en 4 analyseren we de determinanten van de toeloop naar beleggingsfondsen, waarbij we ons concentreren op de rol die beleggingsresultaten uit het verleden daarbij spelen. Hoofdstuk 3 verschaft empirisch bewijs aangaande de dynamiek van de impact van eerdere fondsprestaties op de toeloop van Amerikaanse groeifondsen voor de periode 1991-1999. We identificeren significante niet-lineaire effecten in de vertragingsstructuur van de toeloop-prestatie relatie. In het bijzonder blijkt dat de beleggingsresultaten tijdens het meest recente kwartaal minder belangrijk zijn voor de netto instroom dan de rendementen gedurende de overige drie kwartalen van het jaar, hetgeen suggereert dat sommige investeerders met een vertraging reageren op fondsprestaties. De meeste investeerders lijken alleen de eerste drie jaar van de geschiedenis van fondsprestaties te gebruiken voor hun investeringsbeslissing, hetgeen ongeveer 90 procent van de totale impact van de vroegere prestatie op de toeloop verklaart. Bovendien tonen we aan dat het rendement op systematische risicofactoren ook een invloed heeft op fondstoeloop, naast de impact van risicogecorrigeerde rendementen. Dit houdt in dat sommige beleggers "style timers" zijn, die fondsen kiezen op basis van ruwe prestatie in plaats van risicogecorrigeerde prestatie.

Hoofdstuk 4 levert empirisch bewijs aangaande de impact van verschillende classificatie systemen op de toeloop naar Amerikaanse beleggingsfondsen in 1993-1999. We onderzoeken specifiek de relatie tussen toeloop naar de beleggingsfondsen en hun geordende prestatie binnen drie typen categorieën: fondsen met dezelfde vermelde doelstelling, fondsen met dezelfde Morningstar stijl en fondsen binnen dezelfde beleggingscategorie. We zien dat de prestatievolgorde binnen fondsen met dezelfde Morningstar stijl of in dezelfde beleggingscategorie een sterk positieve invloed heeft op fondstoeloop, bovenop de impact die de prestatievolgorde bij de fondsen met dezelfde vermelde doelstellingscategorie heeft. In feite lijkt een vergelijking van rendementen met die van fondsen binnen dezelfde beleggingscategorie het meest gebruikte referentiekader te zijn voor privé beleggers in US aandelenfondsen en obligatiefondsen. Institutionele beleggers van US aandelenfondsen hechten ongeveer evenveel waarde aan een vergelijking met rendementen van fondsen met dezelfde Morningstar stijl als met rendementen van fondsen met dezelfde vermelde doelstelling. Alleen voor internationale aandelenfondsen is de toeloop-prestatie gevoeligheid het sterkst binnen fondsen met dezelfde vermelde doelstelling. In een gezamenlijk model voor ordinale en kardinale prestatiemaatstaven overschrijdt de impact van het rendement op de fondstoeloop nooit de gecombineerde impact van relatieve prestatie. Wanneer we een categorie-specifieke analyse van het "star spillover effect" uitvoeren (zie Nanda, Wang en Zheng, 2000), zien we dat de aanwezigheid van een topfonds duidelijk voordelig is voor de toeloop naar de andere fondsen uit dezelfde familie. Alleen in het geval van obligatiefondsen "kannibaliseert" de topprestatie van een topfonds de toeloop naar fondsen met dezelfde vermelde doelstelling als het topfonds.

In Hoofdstuk 5 en 6 wordt zowel een speltheoretische als een empirische analyse gepresenteerd van het gedrag van de beheerders van beleggingsfondsen. In Hoofdstuk 5 kijken we naar de in de literatuur uitgevoerde statistische toetsen betreffende het risicogedrag van beheerders van beleggingsfondsen. Busse (2001) bemerkt dat autocorrelatie in dagelijkse fondsrendementen schattingen voor de volatiliteit vertekent, terwijl kruisafhankelijkheid in beleggingsfondsrendementen de onafhankelijkheidsaanname, onderliggend aan de standaard statistische toetsen betreffende risicogedrag, ontkracht. Busse (2001) beweert dat het bewijs ten voordele van de toernooi hypothese, dat wil zeggen de hypothese dat veranderingen in risico gedurende een jaar gerelateerd zijn aan de tussentijdse prestatie van een fonds, verdwijnt wanneer autocorrelatie en kruiscorrelatie in rendementen in beschouwing worden genomen. In Hoofdstuk 5 dragen we bij aan deze discussie door de impact van zowel autocorrelatie als kruiscorrelatie op de toetsen van de toernooi hypotheses mee te nemen vanuit een analytisch standpunt.

Ten eerste geven we een analytische uitdrukking voor de vertekening die ontstaat in schattingen van de volatiliteit (gebaseerd op zowel dagelijkse als maandelijkse gegevens) als gevolg van eerste-orde autocorrelatie effecten in de dagelijkse rendementen. We tonen aan dat toetsen van de toernooi hypothese gebaseerd op maandelijkse gegevens in feite robuuster zijn ten aanzien van autocorrelatie effecten dan toetsen gebaseerd op dagelijkse gegevens. Ten tweede, om de impact van de kruiscorrelatie in rendementen op de toetsen te bekijken, verschaffen we expliciete voorwaarden waaronder de toetsen, zoals gebruikt in de literatuur, de juiste onbetrouwbaarheidsdrempel hebben. Dit toont aan dat een specifieke vorm van cross-sectionele afhankelijkheid in de rendementen van beleggingsfondsen is toegestaan zonder dat de statistische eigenschappen van de toetsen aangetast worden. De cruciale voorwaarde hiervoor is dat diversifieerbare rendementen onafhankelijk tussen fondsen zijn.

In Hoofdstuk 6 bestuderen we incentives betreffende het risicogedrag van beheerders van beleggingsfondsen die relatieve prestatiedoelstellingen hebben (zoals in een toernooi). Ten eerste, in een tweeperioden-model, analyseren we het spel gespeeld door twee risiconeutrale beheerders met relatieve prestatiedoelstellingen. We tonen aan dat de beheerders in de eerste periode hun verwachte rendement maximaliseren, terwijl in de tweede periode de tussentijdse verliezer (i) meer risico neemt dan de tussentijdse winnaar en (ii) dat de mate van risico nemen door de tussentijdse verliezer toeneemt met het verschil in tussentijdse prestaties. Door simulaties te gebruiken demonstreren we dat de keuze van risico van de beheerder in de tweede periode negatief gerelateerd is aan zijn of haar relatieve prestatie over de eerste periode, hetgeen ook het geval is bij meer dan twee concurrerende fondsen. Ten tweede toetsen we op empirische wijze de voorspellingen van het model, gebruikmakend van Amerikaanse aandelenfondsen over de periode 1980-1998. We zien dat de keuze van het systematisch risico in het laatste kwartaal van het jaar negatief gerelateerd is aan de prestatie van het fonds over de eerste drie kwartalen van het jaar. Dit is in overeenstemming met het model.

De analyse van het gedrag van beleggers in en beheerders van beleggingsfondsen in dit proefschrift kan in vele richtingen worden uitgebreid. Een interessant onderwerp van nader onderzoek is de invloed van vroegere rendementen op verscheidene systematische risicofactoren, zoals grootte of momentum, op het gedrag en de strategieën van beleggers en beheerders. Beleggers selecteren hun fondsen wellicht niet alleen op basis van eerdere risicogecorrigeerde prestatie, maar ook op basis van de risico karakteristieken van feitelijke portefeuilles. Zij kiezen mogelijk zo veel mogelijk fondsen met een hoge blootstelling aan "hot" factoren die recentelijk hoge rendementen voortbrachten. Alhoewel Hoofdstuk 3 voorlopig bewijs levert dat sommige beleggers inderdaad rekening houden met de ruwe fondsprestatie, is er een nadere analyse vereist om de afzonderlijke invloeden van verschillende systematische risicofactoren op de beleggingsfondstoeloop te identificeren. Aan de andere kant volgen de fondsbeheerders misschien een vergelijkbare "style-timing" strategie waarbij de blootstelling aan goedpresterende risicofactoren vergroot wordt. Deze strategie zou hen kunnen helpen om het prestatiegat met betrekking tot de fondsen die hun investeringen concentreren in "hot" styles en hogere ruwe rendementen hebben te verkleinen (zie bijvoorbeeld Barberis en Schleifer, 2000). Zulk gedrag zou meer kunnen passen bij beheerders van kleine en volatielere fondsen, aangezien het voor hen makkelijker is significante veranderingen in het investeringsbeleid aan te brengen, alsook bij beheerders van onder de maat presterende fondsen die waarschijnlijk de fondsstrategie gaan veranderen om de fondsprestatie te verbeteren.