ESSAYS ON MUTUAL FUNDS

by

EGEMEN GENC

A THESIS

Presented to the Department of Finance
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2012

DISSERTATION APPROVAL PAGE

Student: Egemen Genc

Title: Essays on Mutual Funds

This thesis has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Finance by:

Dr. Ekkehart Boehmer	Chairperson
Dr. John Chalmers	Member
Dr. Sith Chaousirote	Member
Dr. Jeremy Piger	Outside Member

and

Kimberly Andrews Espy	Vice President for Research & Innovation/Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2012.

DISSERTATION ABSTRACT

Egemen Genc

Doctor of Philosophy

Department of Finance

June 2012

Title: Essays on Mutual Funds

My dissertation consists of two essays on mutual funds. The first essay examines the role of extreme positive returns on future fund flows using maximum style-adjusted daily returns (hereafter MAX) over the previous month. My results suggest that there is a positive and significant relation between MAX and future fund flows. The results are robust to controls for fund performance, fund size, age, turnover, fund fees, volatility, and skewness of fund returns. Of particular interest, this relation exits only in retail funds. Moreover, MAX is persistent from one month to the next, but MAX-based investment strategies are associated with lower risk-adjusted returns than investors could have achieved in otherwise similar funds. Overall, my analysis suggests that mutual fund investors are attracted to maximum style-adjusted daily returns, which is in line with the theoretical argument that investors exhibit a preference for lottery-like payoffs. These investors are successful in achieving a lottery-like return profile, but this strategy is costly in terms of expected returns

The second essay studies the effect of recent and long-term mutual fund performance on future fund flows. I document that investors' response to recent performance depends on average long-term performance. In particular, a recent loser

fund experiences outflows only if its longer-term performance is also poor. Similarly, recent good performance leads to more inflows only if the fund has also good long-run performance. In contrast, investors ignore recent performance if it provides a signal that conflicts with the longer-term signal. This implies that good fund managers with a longer-term focus will find it easier to attract future inflows than managers with a short-term horizon.

CURRICULUM VITAE

NAME OF AUTHOR: Egemen Genc

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Texas A&M University, College Station
Otto-von-Guericke University of Magdeburg, Germany
Bilkent University, Ankara, Turkey

DEGREES AWARDED:

Doctor of Philosophy, Finance, 2012, University of Oregon
Master of Science, Management, 2006, Otto-von-Guericke University
Bachelor of Science, Industrial Engineering, 2000, Bilkent University

AREAS OF SPECIAL INTEREST:

Mutual Funds
Empirical Asset Pricing
Liquidity Risk
Market Efficiency

PROFESSIONAL EXPERIENCE:

Teaching and Research Assistant, Department of Finance, University of Oregon, 2010- 2012

Teaching and Research Assistant, Mays Business School, Texas A&M University, 2007- 2009

Research Assistant, Institute of Technical and Business Information Systems, Otto-von-Guericke University, 2006- 2007

Oracle ERP Consultant and Project Manager, Meteksan Systems Inc., 2000-2005

GRANTS, AWARDS, AND HONORS:

University of Oregon Graduate Assistantship, 2010-2011

Universtiy of Oregon Summer Fellowship, 2010-2011

Flores Fellowship, Texas A&M University, 2007-2010

Depatment of Finance Fellowship, Texas A&M University, 2007-2010

Mays Graduate Assistanship, Texas A&M University, 2007-2010

American Finance Association (AFA) Travel Grant, 2010

DAAD (German Academic Exchange Service) Fellowship, 2005-2007

TEV (Turkish Education Foundation) Fellowship, 2005-2007

Bilkent University Full Scholarship, 1996-2000

Bilkent University, Honor Student List, 1996-2000

ACKNOWLEDGMENTS

I am having a problem finding words, which can adequately express my gratitude to my chair, Dr. Ekkehart Boehmer. He was very supportive and patient in every step of my Ph.D. and inspired me with his guidance and knowledge.

I would like to thank to my committee members, Dr. John Chalmers, Dr. Sith Chaousirote, and Dr. Jeremy Piger for their value inputs in preparing this manuscript. I am especially indebted to Dr. Diane Del Guercio, who relentlessly read my prior versions, and provided insightful suggestions to improve my work.

My PhD journey would not be the same without the emotional and financial support of my parents, Emine and Nazim Genc, my brother, Evren, and my in-laws, Bogumila and Marek Machalski. They cheered me along the way and never once questioned my choice of living and studying so far away from home. Not being able to see them for more than four years has been challenging, but they believe in me and by providing loving and supportive environment, they gave me the necessary strength and motivation to work hard and finish this challenging stage of my life.

Ferhat Akbas, Selcuk Celil, Kyle Tippens, and Will Armstrong became my friends along the way. We encouraged, supported and motivated each other. And mostly, we shared a good deal of laughs together. I will miss them.

Above all, I would like to thank my wife, Agnieszka for being there the whole time. The completion of my PhD would not be possible without her love, invaluable support, and genuine caring. She has been by my side the whole time and endured many hours alone while I worked on my dissertation. She read all my papers, pushed me when

necessary, and was encouraging throughout the whole process. I would have not gotten to where I am today without her.

To my “Minikkusum”

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table Page

Table Page

CHAPTER I

INTRODUCTION

This dissertation consists of two essays, which investigate how investors make choices in mutual funds.

In the first essay, I examine the effect of lottery-like fund returns on future fund flows. One prominent characteristic of individuals is that they often exhibit a preference for lottery-like assets, i.e. assets that have a relatively small probability of a large payoff (Barberis and Huang (2008), Brunnermeier, Gollier, and Parker (2007), Kumar (2009), and Bali, Cakici, and Whitelaw (2011)). Interestingly, even though the participation of individuals in mutual funds has dramatically grown over the last two decades, we have a limited understanding if lottery-like preferences play a role in determining fund flows. In the first essay, using maximum style-adjusted daily returns (MAX) over the previous month as a metric, I show that funds investors are attracted to fund that provide lottery-like returns. Specifically, I show a positive and significant relation between MAX and future fund flows, which is robust to control for other fund characteristics, particularly for skewness and volatility of fund returns. This significant relation only exists in retail funds, which primarily cater to individual investors. This evidence is consistent with the idea that retail investors are more likely to exhibit a preference for lottery-like payoffs (Kumar (2009)). In further tests, I find that MAX in a period is a good predictor of MAX in the subsequent periods, suggesting that if lottery-like features enter investors' utility functions, then choosing such high MAX funds may be optimal. However, I also document that high MAX funds underperform than otherwise similar funds. Thus, fund-selection strategies based on MAX are costly to investors in terms of expected returns.

In the second essay, I investigate the role of short-term and long-term performance on future fund flows. Prior literature finds that investors tend to flock to fund with recent superior performance (Sirri and Tufano (1998), Gruber (1996)). This is consistent with the theoretical model of Berk and Green (2004), who argue that investors infer managerial quality from past performance. Even though prior empirical literature finds a positive relation between flows and recent performance, this does not differentiate between naive return chasing and investor learning. In the second essay, I analyze the sensitivity of flows to recent performance conditional on long-term performance. If some investors learn from past performance, then there is arguably more information about manager skills in returns measured over a longer period compared to returns measured over a shorter period. If so, the longer-term performance history may enhance the confidence with which investors assess the manager’s ability, and affect the sensitivity of flows to recent performance.

My results suggest that investors' decisions are strongly governed by longer-term rather than shorter-term fund performance. In particular, long-term winner funds obtain positive net flows, even though they underperform in the recent quarter. Long-term losers experience outflows despite a superior performance in the recent quarter. My analysis highlights the importance of long-term performance on the shape of the flow-performance relation. I show that favorable long-term performance strengths the sensitivity of flow to recent performance particularly among recent winners. Investors do not ignore the poor recent performance. Instead, they put more weight on the positive long-term performance. This makes it difficult to properly interpret the relation between short-term performance and flows from a model that omits long-term performance.

The findings that both long-term performance and lottery-like return affect future fund flows is consistent with the idea that consumers adopt a multi-attribute model in determining their choices (Fishbein and Azjen (1975), Lanchaster (1966), and Capon, Fitzsimons, and Prince (1996)). As a result, the existence of a preference for lottery-like returns does not necessarily preclude that some investors would positively value long-term performance or vice versa. The value that investors attach to long-term performance and lottery-like returns may differ among different clientele. I show that the preference for lottery-like payoffs only exist in retail funds, which predominantly cater to individual investors. Nevertheless, as long as a sizable portion of investors consider these fund characteristics in their purchase decisions, we can observe the effect of both fund characteristics on fund flows. Accordingly, my tests in each essay show that the effect of MAX or long-term performance does not subsume each other, but survive in the presence of the other.

CHAPTER II

EXTEREME DAILY FUND RETURNS AND FUND FLOWS

Introduction

The ownership of mutual funds by U.S. households has grown dramatically from 6% in 1980 to 44% in 2010. An estimated 90 million individual investors owned mutual funds and held 87% of total mutual fund assets ($ 11.2 trillion) in 2010.[1] Moreover, French (2008) reports that demand for directly held stocks has been decreasing for most of the decade as individuals have increased their reliance on mutual funds. As a result, individuals increasingly determine the demand for mutual funds.

Understanding individuals' choices in mutual funds is important for several reasons. First, investors can indirectly affect stock prices because capital flows into and out of mutual funds determine the demand for stocks in the funds' portfolios. Wermers (2004) finds strong evidence that flow-related additions to existing positions push up stock prices. Similarly, Coval and Stafford (2007) document that the trading activities of funds that experience large inflows or outflows exert price pressure on stocks.[2] Moreover, investors' demand can influence multiple stock prices simultaneously as fund managers tend to either increase or decrease their existing positions in response to fund

[1] See Investment Company Institute Fact Book (2011) (http://www.ici.org/pdf/2011_factbook.pdf).

[2] The authors document that the price pressure is temporary, but the reversal is a slow process that takes three to five quarters. This creates considerable profits for liquidity providers and wealth transfer from the existing shareholders of the funds to the liquidity providers. In addition, Coval and Stafford (2007) argue that price pressure may create some front running incentives, which decrease price efficiency. Consistent with the front running arguments, Chen et. al. (2008) find that some hedge funds profit from flow-related trading of mutual funds. Therefore, even though the price impact is temporary, it is important since there might be a wealth transfer from the existing shareholders of the funds to the liquidity providers or front runners. On the other hand, discretionary trading by mutual funds does not cause reversal (see Coval and Stafford (2007), and Khan, Kogan, and Serafaim (2011)), and consequently no wealth transfer occurs. From this perspective, flows caused by investors' preferences affect both prices and the wealth of fund investors.

flows rather than create new stock positions (Pollet and Wilson (2008)). Considering that mutual funds are one of the largest groups of investors in U.S. companies, holding 28% of the outstanding stocks, the economic impact of flow-induced trading can indeed be substantial.

Second, since fund managers generally receive a fixed percentage of assets under management as a fee, fund flows are the mechanism by which investors influence managerial incentives. Consistent with this argument, the literature shows that investors' reactions to different fund characteristics drive managers' behavior.[3] Therefore, understanding the role of investors' preferences as a determinant of fund flows is crucial in order to decipher managerial actions. Third, financial researchers and practitioners generally advocate investing into well-diversified funds with low turnover and low expense ratios as a smart investment strategy. Yet, investors seem to ignore some of these basic principles, and thus additional evidence on how they choose among funds is useful to understand investors' decision making process per se.

One prominent characteristic of individuals is that they often exhibit a preference for lottery-like assets, i.e. assets that have a relatively small probability of a large payoff. Barberis and Huang (2008) and Brunnermeier, Gollier, and Parker (2007) develop models which provide theoretical motivation for this preference. Kumar (2009) shows that certain groups of individuals are inclined to hold lottery-type stocks, which subsequently underperform on average. Bali, Cakici, and Whitelaw (2011) suggest that stocks with extreme positive returns attract investors.

[3] See Chevalier and Ellison (1997), Brown, Harlow, and Starks (1996), Cooper, Gulen, and Rau (2005), Sensoy (2009).

In this paper, I investigate the role of extreme positive returns on future fund flows using maximum style-adjusted daily returns (hereafter MAX) over the previous month. My results suggest that there is a positive and significant relation between MAX and future fund flows. The difference in flow ratios, defined as the growth rate of total net new money, between funds in the highest and lowest MAX deciles in the univariate portfolio analysis is around 0.84% per month. This finding is robust to a battery of bivariate and multivariate tests, which control for various fund characteristics including performance, total volatility, total skewness, idiosyncratic volatility, idiosyncratic skewness, size, age, turnover, marketing costs (i.e. 12b-1), expenses, and load fees. In particular, a 1% increases in MAX increases the flow between 0.11% and 0.24% per month in multivariate tests. In addition, I obtain similar results when using the average of the two, three, four, and five highest daily returns within the month as alternative proxies for extreme returns. Finally, I consider a measure defined as the *average* of single day MAX over the past two, three, six, twelve, and eighteen months. The effect of MAX gets even stronger when I average over longer periods. Overall, the evidence suggests that fund investors may be willing to direct more flows to funds that have recently exhibited large positive returns.[4]

One interpretation of my findings is consistent with the theoretical models of Barberis and Huang (2008), and Brunnermeier, Gollier, and Parker (2007). Using the cumulative prospect theory of Tversky and Kahneman (1992), Barberis and Huang

[4] In order for investors to be attracted to MAX, at least some investors should have access to information on daily returns of mutual funds. Many financial sites (i.e. WSJ Fund Screener, Bloomberg U.S. Fund Rankings) report daily returns, daily rankings of funds within each investment style as well as historical highest and lowest prices and returns within a month. According to the ICI Fact book 2011, 82% of fund investors use online sources for financial purposes, mostly to obtain investment information. Therefore, besides other possible venues that investors might utilize to reach daily returns of mutual funds, there are already publicly available sources that investors can access with minimal effort.

(2008) argue that investors overweight the tails of asset returns. Brunnermeier, Gollier, and Parker (2007) develop a model in which investors choose to increase their beliefs about probabilities of positive payoffs in order to maximize their current utility. Both models predict that investors overvalue the assets that have a small probability of a large positive return

In order to better understand the flow-MAX relation, I further categorize funds as retail vs. non retail funds and perform the analysis within each group. I find that the flow-MAX relation exists only among funds catering primarily to retail investors. Specifically, when I conduct multivariate tests within retail funds vs. non-retail funds, the coefficient of MAX is positive and highly statistically significant only in retail funds, ranging from 0.133 (t-statistics: 4.26) to 0.241 (t-statistics: 5.45) depending on the specification. However, there is no relation between MAX and flows in non-retail funds. In other words, preference to extreme returns exists mainly in retail investors, while the relation between flow and other fund characteristics such as performance and fees holds for both groups. This evidence is consistent with existing results that retail investors are more likely to exhibit a preference for extreme payoff states (Kumar (2009)).

To the best of my knowledge, this is the first study that documents the importance of preferences for lottery-like features in the distribution of mutual fund returns on mutual fund flows. Recently, using a sample of individual investors from a brokerage house between 1991 and 1996, Bailey, Kumar, and Ng (2011) examine whether a combination of behavioral factors jointly affects the use of mutual funds in individuals' portfolios. They argue that a certain class of investors is affected by behavioral biases.

These investors participate less in mutual funds[5], but when they do, they choose funds with relatively high expense ratios and turnover. Bailey, Kumar, and Ng (2011) do not examine, however, the relation between investor decisions and fund flows, or the relationship between flows and extreme payoffs. For fund flows to be affected by the preference of some investors, collective decisions of these investors must lead to an overlay increase or decrease in fund flows, which cannot be corrected by arbitrageurs or sophisticated investors. Bailey, Kumar, and Ng (2011) argue that there are “smart “ investors that are not affected by behavioral biases and they make investment decisions, which are opposite to the decisions made by the investors with behavioral biases.

Therefore, studying fund flows can shed light on the consequences of investors' decisions at the fund level and reveal an important link between investor decision making and asset prices as well as investor decision making and managerial incentives. Additionally, a sample of a brokerage house may exhibit different characteristics than the general population. For example, Bilias, Georgarakos, and Haliassos (2009) find stark contrast between the trading behaviors of the investors with brokerage accounts and the general population, and argue that the brokerage sample may represent a nonrandom sample.[6] Hence, even though the sample used in Bailey, Kumar, and Ng (2011) provide

[5] Participation is defined as a dummy variable that is equal to one if an investor invests in mutual funds at least once during the sample period.

[6]The authors use Survey of Consumer Finances (SCF) data that differentiates between investors that hold brokerage accounts and the ones that do not hold and also include many characteristics of these investors. They find that 20% of overall population holds brokerage accounts and they significantly trade more than the general population. Additionally, the investors who own such accounts tend to invest only a small fraction of their financial wealth. The median brokerage account as a share of household financial wealth for brokerage account owners is around 10% or less, while it is 3.5 % as a share of net total wealth. The authors also argue that brokerage account owners may represent a nonrandom sample as the probability of having a brokerage account depends on the characteristics of the investors such as education, willingness to take above average financial, intention to leave a bequest, children etc. As a result, Bilias, Georgarakos, and Haliassos (2009) concluded "…To the extent trading in brokerage accounts induces volatility,

unique data on individuals, we need additional research to understand whether the effect of investors' decisions show up at the fund level.

This paper also complements many papers that examine various dimensions of individuals' fund choices and the determinants of mutual fund flows. Barber, Odean, and Zheng (2005) find that investors are more sensitive to in-your-face fees such as loads than to operating expenses, which are less visible. Sirri and Tufano (1998) show that funds receiving more media attention attract more capital. In a similar vein, Jain and Wu (2000) and Gallaher, Kaniel, and Starks (2008) provide evidence that funds with higher marketing efforts receive more flows, while Del Guercio and Tkac (2008) show the response of flows to changes in ratings. Different from the previous literature, my analysis links investor preferences for lottery-like features to fund flows and thus, to determinants of asset prices.

There are empirical studies that support the theoretical models that justify the preference for lottery-like payoffs (Barberis and Huang (2008), Brunnermeier, Gollier, and Parker (2007)). In a laboratory experiment with a sample of 148 students and 131 experienced executives, Åstebro, Mata, and Santos-Pinto (2009) find that subjects make skew seeking choices, especially when the upside gain is high. Similarly, Boyer, Mitton, and Vorkink (2010) and Conrad, Dittmar, and Ghysels (2009) find that investors are attracted to stocks with high expected skewness. Boyer and Vorkink (2011) show a similar preference in stock options. Most recently, Bali, Cakici, and Whitelaw (2011) use the maximum daily return over a month and document a significant negative relation

overtrading may be relevant for asset pricing, but it seems dubious that it is a major importance for household finances".

between these extreme positive returns and expected stock returns (i.e. the monthly average returns in the subsequent month). They interpret this relation as evidence that investors are willing to pay more for stocks with lottery-like returns, which yield lower returns in the future. The evidence suggests that investors' preferences to high payoffs may play a significant role in determining asset returns.

However, despite the important implications of fund flows on asset prices and managerial incentives, we have limited understanding of how lottery-like returns affect product choices in the mutual fund market. First, investors may invest in mutual funds for liquidity and diversification reasons, which they may not able to accomplish by investing in individual stocks, or they may prefer mutual funds as a passive form of investing rather than a more active form of investing in stocks. Therefore, while investors may present a tendency to buy stocks with lottery-like returns when they invest in the stock market, it is not clear whether they will also present a similar behavior when they choose professional managers to invest their money. Second, even if some investors exhibit this behavior, it is an empirical question whether fund flows will be affected by these preferences. Third, other studies, infer investors' preferences from changes in prices. For example, Bali Cakici (2011) find that stocks with extreme daily returns in a month are related to negative expected returns which is one prediction of the theoretical models (Barberis and Huang (2008)). While the link between extreme daily returns and expected returns is the excess demand for the lottery-like stocks, they do not directly show this link, but infer investor’s preference from low future returns. On the other hand, fund flows measure the money in and out of the funds, which is determined by the investors’ choices, and hence mutual funds provide a unique opportunity to observe investors' preferences via fund flows (Cooper, Gulen, and Rau (2005)).

My tests do not allow inferences about whether MAX-based fund selection strategies are optimal or not. If MAX is a predictor of risk-adjusted future performance, then selecting funds based on MAX may be simply an easy way for investors to select well-managed funds. But even if MAX-based fund investments underperform, their selection could be rational. For example, if lottery-like features enter investors' utility functions, then choosing such funds may be optimal if MAX predicts future MAX. In other words, if lottery-like features are persistent and valuable to investors, then MAX-based flows can be rational.

I find evidence that MAX-based fund selections underperform, but I also document that MAX is robustly persistent over time. Specifically, my analysis suggests that funds in the top decile with respect to MAX are in the top decile (in one of the top three deciles) in the subsequent month with a probability of 48% (79%). Multivariate fund-level cross-sectional regressions suggest that MAX demonstrates significant persistence with a positive coefficient of 0.612 after controlling for various fund characteristics including size, age, volatility, skewness, return, flow, expenses, and load fees. Regarding performance, I show that funds with a 1% higher level of MAX are associated with 10 bps lower future raw returns per month and 8 bps lower future risk adjusted returns per month. Thus, fund-selection strategies based on MAX are costly to investors in terms of expected returns.

There are a handful of papers that discuss whether investors are smart in their capital allocations among funds. Gruber (1996) and Zheng (1999) show that fund inflows are associated with higher future performance and fund flows appear to be "smart money". However, Frazzini and Lamont (2008) argue that this effect is, at best, confined to short horizons and investors make suboptimal decisions in the long run. This conclusion is reinforced by Cooper, Gulen, and Rau (2005), who show that loser funds attract inflows by simply changing their names to a hot style without actually changing their style. Moreover, Elton, Gruber, and Busse (2004) find that, even in the simplest investment products such as index funds, naive rules in selecting funds (i.e. low expenses and higher past returns) outperform the returns that investors actually earn. This paper also contributes to this literature by showing that investors are successful in achieving lottery-like payoffs in the future, but this success comes at the cost of lower expected returns.

The rest of this chapter is organized as follows. In section II, I describe my sample and my main variable (MAX), and motivate the use of this variable as a proxy for lottery-like preferences. In section III, I analyze the effect of MAX on fund flows using portfolio analysis and multivariate regressions. Section IV interprets the results. Section V reports the results from robustness tests. Section VI offers concluding remarks.

Data

Sample

Daily and monthly fund returns, and other fund characteristics are obtained from the Center for Research in Security Prices (CRSP) Survivor-Bias Free U.S. Mutual Fund Database. The sample covers domestic equity funds from October 1998, the first month CRSP starts reporting daily fund returns until December 2010. I select funds with the following Lipper investment categories[7]: *(i) Small-Cap Growth ('SCGE'), (ii) Small-Cap Core ('SCCE') (iii) Small-Cap Value ('SCVE'), (iv) Mid-Cap Growth ('MCGE'), (v) Mid-Cap Core ('MCCE') (vi) Mid-Cap Value ('MCVE'), (vii) Multi-Cap Growth ('MLGE'), (viii) Multi-Cap Core ('MLCE'), (ix) Multi-Cap Value ('MLVE'), (x) Large-Cap Growth ('LCGE'), (xi) Large-Cap Core ('LCCE'), (xii) Large-Cap Value ('LCVE').* This filter removes bond funds, balanced funds, international funds, and sector funds from the sample. To further remove index funds, I use the "*index fund flag*" field available in CRSP as well as fund names.[8] In order to account for a potential incubation bias, I eliminate funds that are less than two years old and observations with a missing fund name in CRSP.[9] From the remaining sample, I further exclude funds with zero expenses since observations with zero expenses are most likely to indicate missing information (Barber, Odean, and Zheng (2005), and Gil-Bazo and Ruiz-Verdΰ (2009)). The final sample contains 719,538 fund-month observations from 11695 distinct funds and 12

[7] CRPS provides three different codes to identify fund investment styles: Wiesenberger codes, Strategic Insights (SI) codes, and Lipper codes. These codes cover three different time periods: The Weisenberger codes are available before 1992; the Strategic Insights codes are available between 1992 and 1998; and the Lipper codes exist after 1999. Since my sample starts from October 1998, I only use Lipper codes. For the remaining four months in 1998, I select the funds that are in my sample after 1999. If a Lipper code is missing in a particular year, but is available for that fund in a later or earlier year, then I keep that observation and fill the code with its earlier or later classification. Moreover, if a fund's Lipper code is not same in each year, then the fund is assigned to its most frequently classified category as in Pastor and Stambaugh (2002).

[8] A value "D" in the "index fund flag" indicates a pure index fund. However, this flag is available only after June 2008. Hence, I also make a name search to identify index funds. Following Gil-Bazo and Ruiz-Verdΰ (2009), I code a fund as an index fund if its name contains any of the following strings: "Index", "Idx", "Ix", "Indx", "NASDAQ", "Nasdaq", "Dow", "Mkt", "DJ", "S&P 500", "BARRA".

[9] Mutual fund incubation is a strategy in which fund families start multiple new funds and open some of them to the public at the end of an evaluation period, while the others are shut down before investors become aware of them. This strategy creates an upward bias in fund returns. Evans (2010) suggests that an age filter eliminates this bias.

investment styles as defined by Lipper investment categories.[10]

Variable Definitions

I define the monthly net flow into a fund as:

$$\text{Flow}_{i,t} = \frac{\text{TNA}_{i,t} - \text{TNA}_{i,t-1} * (1 + \text{R}_{i,t})}{\text{TNA}_{i,t-1}} \tag{1.1}$$

where $\text{R}_{i,t}$ is the monthly return of fund i during month t and $\text{TNA}_{i,t}$ is the total net asset value of fund i at the end of month t. This definition reflects the growth rate of a fund due to new investments and assumes that all new investments occur at the end of the month. Elton, Gruber, and Blake (2001) show that the dates of fund mergers often differ from actual merger dates and this inaccuracy introduces a large number of errors in fund returns. Even though this bias in fund returns does not show a systematic pattern, it may lead to extreme values of flows. Therefore, following Huang, Wei, and Yan (2007), I filter out the top and bottom 2.5% of tails of the flow data in order to mitigate the impact of potential outliers around mergers.[11]

The main variable of interest, MAX_t, is defined as the style-adjusted maximum daily return within month t. Style-adjusted daily returns are calculated by subtracting the daily returns of a fund from the average daily returns of all funds with the same style (Teo and Woo (2001)). Since mutual funds are often confined to trading stocks within their style, returns of the funds within the same style have high cross-sectional correlations. Hence, funds with superior returns are more likely to belong to above-

[10] I treat each share class as a distinct fund. For the purpose of the paper, it is important to conduct tests at the share class level because different share classes may cater to different investors (Nanda, Wang, and Zheng (2009)). In particular, funds may offer both institutional and retail classes. Investors in these classes behave differently regarding the lottery preferences (see section IV). Nevertheless, analysis at the fund level gives qualitatively similar results.

[11] Alternatively, I filter out the top and bottom of 1% tails of the flow data. Results are similar.

average performing styles. Style-adjusted returns control for the time-varying style effect, and mitigate the concern of categorizing funds as high MAX funds because of the popularity of the style they belong to.[12] Moreover, given the frequent references to fund categories in media and other financial resources, it is mostly likely that individuals use style information in their investment decision in order to simply information processing (Mullainathan (2001)). Empirically, there is considerable evidence that investors make their decision based on fund and style returns (see Barberis and Shleifer (2003), Pomorski (2004)). Therefore, it is particularly important to adjust for style returns in order to differentiate style attraction from the effect that I want to capture.

MAX is a well-motivated measure from a theoretical perspective and it differs conceptually and empirically from other measures of return dispersion or asymmetry such as skewness or volatility. In Barberis and Huang (2008), positive extreme states are the mechanism through which investors maximize their value function, and consequently overvalue assets that provide these states. Similarly, in Brunnermeier, Gollier, and Parker (2007), the pricing effects are driven by extreme positive states not by skewness per se. Moreover, in both models, investors' attraction to assets comes from the anticipation of extreme payoffs in the future, requiring that a measure should be persistent to be used for both models' predictions. My tests show that MAX satisfies this criterion.[13] MAX is

[12] I also use value-weighted and median returns in each style to make the adjustment (Gaspar, Massa, and Matos (2006), and Sialm and Tham (2011)). Results are similar and available upon request. I also use unadjusted MAX as a sorting variable. Even though high MAX funds predominantly belong to small-growth funds, making harder to differentiate between MAX effect and style effect, I still observe that high MAX funds have higher flows than low MAX funds.

[13] In section IV A, I show that MAX in a given period strongly predicts MAX in the subsequent periods. However, among others, Singleton and Wingender (1986) show that skewness in a given time period may not be a good predictor of skewness in the subsequent period. This point is also valid for idiosyncratic skewness (Harvey and Siddique (2000), and Boyer, Mitton, and Vorkink (2010)).

based on the right tail of a return distribution and may better capture the notion that investors primarily judge on lottery-like return profile based on the tail events as opposed to the entire distribution used in skewness (Barberis and Huang (2008), Bali, Cakici, and Whitelaw (2011)).

From a practical standpoint, MAX is also a simple measure to understand and is easily accessible. Many financial sites (i.e. WSJ Fund Screener, Bloomberg U.S. Fund Rankings) report daily returns, the daily rankings of funds within each investment style as well as the historical highest and lowest daily prices and returns in a given month. Presumably, investors can grasp MAX more easily than other measures such as skewness, which requires a calculation of third moments and is generally not supplied in financial resources.

Size is defined as the total net assets (TNA) of a fund. Age is the number of years since the inception date of a fund.[14] Expense ratio is the fund total operating expenses expressed as a percentage of the fund's average net assets. 12b-1 fees are fees paid out of fund assets to cover marketing, advertising, and distribution services. Non12b-1 fees are the expense ratio net of 12b-1 fees and are used to proxy non-marketing expenses.[15] Front load (rear load) measured as a percentage of fund assets is the fee investors pay when they buy (redeem) fund shares. Return is the monthly fund return net of expenses. Four-factor alpha is the intercept from the regression of monthly fund excess returns on

---

[14] I use "*first_offer_date*" as a proxy for the fund inception date.

[15] Barber, Odean, and Zheng (2005) show that flows seem to react differently to 12b-1 fees and Non12b-1 fees. 12b-1 fees are mainly used for brokers' compensation or for advertising, both of which may increase the visibility of the funds. On the other hand, Non12b-1 fees cover portfolio management and administrative expenses payable to the fund's investment advisor as well as custodial, legal, and accounting expenses.

Carhart's (1997) four factors in the preceding 60 months. Volatility of a fund in month t is the standard deviation of daily returns within month t, and skewness of a fund in month t is calculated from the daily returns over the previous 12 months.

Table 1.1 presents summary statistics of my sample. For each key statistics, I first compute a cross-sectional mean in each month within each investment style, and then report the time series average of these cross-sectional means. In a given month the sample includes, on average, 4944 funds with average total net assets (TNA) of $490.42 million. Large-cap core funds are the largest group with an average number of 744 funds, while mid-cap value is the smallest groups with an average number of 207 funds. On average, funds charge similar fees, even though non-marketing fees for small-cap funds are slightly higher. Growth oriented funds tend to have higher turnover, which indicates a shorter-term strategy to pursue supposedly superior returns. Their trading strategy also leads to more volatile returns for these funds. Small and mid-cap funds earn higher returns net of expenses (0.88% and 0.84% on average) than other funds. However, there is no discernible difference in the risk-adjusted returns (four factor alphas) among funds.

Characteristics of MAX Sorted Portfolios

Funds that deliver MAX might have certain characteristics different from other funds. To get a picture of the composition of high MAX portfolios, each month I rank all funds into deciles on the basis of MAX and calculate the cross-sectional mean of fund characteristics including MAX, size, age, style-adjusted monthly returns, fund fees, turnover, volatility, and skewness. All variables are calculated as in Table 1.1. Panel A of Table 1.2 shows the time series averages of the cross-sectional means for these variables. In addition, Panel B Panel B shows the correlations among the same variables.

Table 1.1. Summary Statistics

This table reports the time-series averages of monthly cross-sectional mean values for various fund characteristics from 1998/9 to 2010/12. Mutual funds are divided in categories according to their Lipper classifications. N is the average number of funds within each category. Size is the total net assets of funds (TNA). Age is the number of years since the inception date. Expense ratio is total annual management and administrative expenses. 12b-1 fee is the marketing, advertising and distribution expenses, while Non12b-1 fee- measured as the difference between expense ratio and 12b-1 fee- is a proxy for non-marketing fees. Front (rear) load is the fee that investors pay when they buy (redeem) fund shares. Return is the monthly style-adjusted return while the four-factor alpha is the intercept from the regression of monthly fund excess returns on Carhart's (1997) four factors in the preceding 60 months. Flow is the percentage change in TNA adjusted for investment returns. Turnover is defined as the minimum of purchases and sales over average TNA. Volatility is the standard deviation of returns in a month and skewness is calculated using daily returns in each year. % of load funds is the number of funds that charge front and/or rear loads.

Category	N	TNA	Age	Expense Ratio	Non 12b-1	12b-1	Front Load	Rear Load	Return	Four-factor alpha	MAX	Flow	Turnover	Volatility	Skewness
LCCE	744	672.03	10.63	1.34%	0.97%	0.37%	0.99%	0.90%	0.36%	-0.11%	0.52%	0.08%	82.76%	1.19%	2.56%
LCGE	630	547.10	9.16	1.43%	1.04%	0.40%	0.93%	1.00%	0.39%	-0.09%	0.64%	0.48%	96.60%	1.31%	5.38%
LCVE	388	852.24	9.91	1.36%	0.94%	0.42%	1.08%	1.00%	0.46%	-0.10%	0.50%	0.37%	65.64%	1.13%	-0.56%
MCCE	209	396.39	8.03	1.64%	1.29%	0.35%	0.95%	0.99%	0.86%	0.03%	0.82%	0.91%	106.26%	1.28%	-7.86%
MCGE	446	264.40	8.83	1.56%	1.16%	0.40%	0.97%	1.00%	0.80%	-0.06%	0.85%	0.62%	132.28%	1.49%	-4.25%
MCVE	207	397.81	7.72	1.50%	1.13%	0.37%	0.98%	0.91%	0.86%	0.02%	0.72%	1.02%	87.54%	1.18%	-9.15%
MLCE	531	383.38	7.85	1.36%	1.00%	0.36%	0.91%	0.87%	0.53%	-0.06%	0.75%	0.68%	84.54%	1.17%	-3.20%
MLGE	352	950.20	8.92	1.57%	1.15%	0.42%	1.02%	1.03%	0.58%	-0.01%	0.97%	0.64%	126.93%	1.44%	0.09%
MLVE	363	495.98	8.93	1.34%	0.97%	0.36%	0.92%	0.90%	0.54%	-0.08%	0.67%	0.47%	69.96%	1.13%	-3.78%
SCCE	435	228.13	7.45	1.64%	1.32%	0.31%	0.77%	0.94%	0.90%	-0.08%	0.78%	0.73%	92.69%	1.30%	-7.53%
SCGE	425	157.50	7.58	1.66%	1.30%	0.36%	0.88%	1.02%	0.85%	-0.13%	0.87%	0.50%	133.82%	1.51%	-7.42 %
SCVE	214	178.27	7.32	1.59%	1.22%	0.37%	0.94%	1.00%	0.89%	-0.11%	0.71%	0.66%	70.70%	1.24%	-5.77%
All funds	4944	490.42	8.76	1.47%	1.10%	0.38%	0.94%	0.96%	0.60%	-0.07%	0.71%	0.52%	92.69%	1.28%	-0.43%

Table 1.2. Characteristics of Funds Sorted by MAX

Each month t from 1998/9 to 2010/12, funds are sorted into deciles based on MAX in month t. Decile 1(10) is the group of funds with the lowest (highest) MAX. For each decile. Panel A reports the time series averages of the cross-sectional mean values for various fund characteristics in the same month. Size is the total net assets of funds (TNA). Age is the number of years since the inception date. Expense ratio is total annual management and administrative expenses. 12b-1 fee is the marketing, advertising and distribution expenses, while Non12b-1 fee- measured as the difference between expense ratio and 12b-1 fee- is a proxy for non-marketing fees. Front (rear) load is the fee that investors pay when they buy (redeem) fund shares. Return is the monthly style-adjusted return. 12MRet is the cumulative twelve months returns Volatility is the standard deviation of daily returns in a month and skewness is calculated using daily returns over the previous 12 months. Panel B reports time series averages of cross-sectional correlation for the same fund characteristics.

Panel A: Fund Characteristics

	MAX	Size	Age	Turnover	Expense Ratio	Front Load	Rear Load	12b-1 Fee	Flow	Return	FF4 alpha	12MRet	Volatility	Min	Skewness
Low MAX	0.23%	766.15	8.99	82.44%	1.30%	0.93%	0.94%	0.60%	0.28%	-0.03%	-0.11%	3.37%	1.17%	-0.37%	-0.07%
2	0.34%	584.09	9.06	83.48%	1.35%	0.97%	0.96%	0.61%	0.36%	0.12%	-0.11%	4.42%	1.19%	-0.45%	-0.78%
3	0.41%	540.86	8.98	86.00%	1.39%	0.96%	0.98%	0.60%	0.42%	0.26%	-0.10%	4.95%	1.21%	-0.50%	-1.24%
4	0.48%	487.31	8.83	88.65%	1.41%	0.97%	0.98%	0.60%	0.49%	0.37%	-0.10%	5.54%	1.23%	-0.55%	-1.60%
5	0.55%	491.70	8.77	91.19%	1.44%	0.97%	0.99%	0.61%	0.54%	0.49%	-0.09%	5.89%	1.24%	-0.60%	-1.93%
6	0.63%	458.33	8.74	93.16%	1.46%	0.97%	0.99%	0.61%	0.51%	0.59%	-0.08%	6.09%	1.27%	-0.65%	-1.93%
7	0.72%	443.53	8.65	95.77%	1.49%	0.96%	0.98%	0.60%	0.56%	0.69%	-0.07%	6.35%	1.29%	-0.72%	-2.25%
8	0.85%	422.70	8.58	99.86%	1.51%	0.95%	0.99%	0.60%	0.56%	0.85%	-0.05%	6.96%	1.32%	-0.80%	-2.75%
9	1.06%	410.30	8.45	105.61%	1.54%	0.93%	0.96%	0.59%	0.70%	1.13%	-0.03%	7.79%	1.35%	-0.93%	-3.31%
High MAX	1.84%	349.42	8.57	131.31%	1.83%	0.86%	0.87%	0.57%	0.86%	1.55%	0.00%	8.28%	1.57%	-1.46%	-2.70%

Table 1.2. (Continued)

Panel B: Correlation among fund characteristics

	MAX	Return	12Mret	FF4 alpha	Size	Age	Turnover	Exp. Ratio	12b-1	Front Load	Rear Load	Volatility	Skewness
MAX	1												
Return	0.191	1											
12MRet	0.039	0.230	1										
FF4 alpha	0.035	0.035	0.311	1									
Size	-0.030	0.007	0.019	0.065	1								
Age	-0.011	0.000	-0.016	-0.035	0.311	1							
Turnover	0.144	-0.005	-0.025	-0.080	-0.064	-0.052	1						
Exp. Ratio	0.245	-0.060	-0.104	-0.181	-0.107	-0.057	0.108	1					
12b-1	-0.029	-0.019	-0.049	-0.105	-0.107	-0.179	0.003	0.671	1				
Front Load	-0.014	0.000	-0.011	0.003	0.062	0.166	-0.014	-0.041	-0.528	1			
Rear Load	-0.021	-0.016	-0.042	-0.078	-0.057	-0.076	-0.018	0.302	0.591	-0.206	1		
Volatility	0.449	-0.015	0.001	-0.076	-0.041	-0.032	0.197	0.167	-0.011	-0.015	-0.009	1	
Skewness	-0.001	-0.001	-0.186	-0.039	0.006	0.021	-0.067	0.014	0.004	0.006	-0.002	0.005	1

By construction, MAX increases from 0.23% in decile 1 (low MAX) to 1.84% in decile 10 (high MAX) indicating a sizeable variation in MAX across deciles. The mean style-adjusted return for high MAX (decile 10) funds is 1.55%, while it is -0.33% for low for low MAX (decile 1) funds. Similarly, past twelve month returns and four-factor alpha estimated from previous 60 months also greater in high MAX funds. The contemporaneous correlation between style-adjusted returns and MAX is 19%, while the correlation between twelve month returns and four-factor alpha is only around 4%. Total net assets of funds decrease as MAX increases across deciles. That is, high MAX portfolios are dominated by relatively smaller funds even though the absolute numbers are difficult to interpret due to growth of the mutual fund industry over time.[16] High MAX funds tend to trade much more, and consequently charge higher operating fees compared to low MAX funds. Since volatility is calculated from daily returns in the same month, it is not surprising that high MAX funds have higher volatilities. The correlation between volatility and MAX is 45%.

Fund returns in my sample are negatively skewed on average and skewness of all MAX is also negative. When skewness is measured over the previous 12 months as in Table 1.2, high MAX funds are more negatively skewed than low MAX portfolios and consistently, skewness of all MAX portfolios is negative and are more negative for high MAX portfolios. However, when skewness is measured over shorter horizons, then high MAX portfolios become more positively skewed. For instance, when skewness is calculated using daily returns in the same month when MAX is measured, then high

[16] I replicate the single sort analysis in Table 1.2 by excluding the funds in the lowest quintile according to size. The results get even stronger after this filter. Moreover, I construct double sorted portfolios on MAX and size. The results are documented in Table 1.5 and show that the effect of MAX is stronger in bigger funds.

MAX portfolios have an average skewness of 5.01%, while low MAX portfolios have an average skewness of -0.97%. Moreover, interestingly, there is no significant correlation between skewness and MAX, suggesting that MAX is hardly a proxy for skewness, or vice versa.

Flow-MAX Relation

Portfolio Analysis: Univariate and Bivariate Sorts

In this section, I use both univariate and bivariate (sequential) sorts to investigate the effect of MAX on future fund flows. In univariate sorts, I construct portfolios based on MAX and analyze future flows. In bivariate sorts, I examine the relation between MAX and future flows after controlling fund characteristics including style-adjusted returns, size, expense ratio, turnover, volatility, and skewness.

I begin the analysis with univariate sorts. Each month t, I rank all funds according to the style-adjusted maximum daily return (MAX_t) within month t and assign them into deciles. These decile portfolios are held for a month after the portfolio formation. High MAX (decile 10) represents highest MAX decile whereas low MAX (decile 1) represents lowest MAX decile. Table 1.3 reports equally weighted monthly portfolio flows ($Flow_{t+1}$) for each decile in month t+1.

According to the results, the average flow difference between high MAX and low MAX decile is 0.84% per month with a corresponding Newey and West (1987) t-statistics of 4.83. [17] The average dollar value of this flow difference (measured as the flow difference times the average size) is around $49 million per year. This difference is also

[17] I also calculate TNA-weighted flows. The difference in flows between high MAX and low MAX is 0.81% and is statistically and economically significant.

economically significant especially given the evidence of a competitive market for mutual funds after 1998 (Wahal and Wang (2011)).

Table 1.3. Average Monthly Flow of Funds Sorted by MAX(N)

Each month t from 1998/9 to 2010/12, funds are sorted into deciles based on the average of N (N=1, 2, 3, 4, 5) highest daily style-adjusted returns (MAX (N)) in month t. These decile portfolios are held for a month after portfolio formation. This table reports equally weighted monthly portfolio flows for each decile in month t+1. Decile 1(10) is the group of funds with the lowest (highest) maximum multi-day returns. Flow is the growth rate of assets under management due to new investments. The last row (HM-LM) presents the difference of flow ratios between the High MAX (decile 10) and Low MAX (decile 1) groups. Newey and West (1987) adjusted t-statistics (up to 6 lags) are reported in parenthesis. All values except the t-statistics are given in percentage terms

Deciles	MAX	MAX (2)	MAX (3)	MAX (4)	MAX (5)
Low MAX	0.29	0.28	0.27	0.26	0.26
2	0.40	0.40	0.40	0.39	0.39
3	0.40	0.39	0.37	0.35	0.34
4	0.51	0.47	0.49	0.48	0.47
5	0.55	0.54	0.51	0.50	0.48
6	0.57	0.58	0.59	0.61	0.62
7	0.62	0.61	0.62	0.59	0.59
8	0.64	0.65	0.64	0.67	0.69
9	0.79	0.77	0.79	0.79	0.79
High MAX	1.13	1.20	1.23	1.25	1.27
HM-LM	0.84	0.92	0.96	0.99	1.01
	(4.83)	***(5.05)***	***(5.25)***	***(5.21)***	***(5.19)***

Conditioning on a single day in defining MAX is a simple and intuitive way to proxy extreme positive returns. However, in order to ensure the robustness of the results, I also use alternative definitions of MAX. First, I rank all funds into deciles by the average of two, three, four, and five highest style-adjusted returns (MAX $[N]_t$, N=2, 3, 4,

5) within month t. Again, decile portfolios are held for a month and the average flows in month t+1 ($Flow_{t+1}$) for each portfolio are calculated. The results are presented in columns 2 to 5 of Table 1.4. I observe a similar pattern in future flows when I use these alternative definitions of MAX. In all cases, the flow difference between high MAX [N] (decile 10) and low MAX [N] (decile 1) portfolios is significant at 1% level, and in fact increases as I average over more days.

Second, I use longer horizons to calculate MAX. Specifically, each month t, I calculate the single style-adjusted maximum daily return as well as the average of two, three, four, and five highest style adjusted returns ($MAX[N]_{t,k}$, N=1,2,3,4,5, K=3,6,12) over the previous three, six, and twelve months. I rank funds into deciles according to these alternative definitions of MAX. Again, decile portfolios are held for a month and the average flows in month t+1 ($Flow_{t+1}$) for each portfolio are calculated. The results are presented in Panel A, B, C of Table 1.4. For all horizons, the flow difference between highest and lowest $MAX[N]_k$ deciles ranges between 0.66% and 0.73% per months and is highly statistically significant. Results get even stronger as I average over more days over the past three, six, and twelve months. Thus, not only the very recent extreme returns but also extreme returns observed over a longer period may affect investors' decisions. As a final check, I use the average of MAX (i.e. maximum style-adjusted daily return in a month) over the previous two, three, six, twelve, and eighteen months ($AMAX_N$ N=2, 3, 6, 12, 18). Panel D of Table 1.4 reports the average flows in month t+1 of decile portfolios sorted on these measures. The difference in average flows between the highest and lowest $AMAX_2$ (the average of MAX over two months) is 0.87% and increase monotonically to 1.02% for $AMAX_{18}$ (the average of MAX over eighteen months). This

Table 1.4. Average Monthly Flow of Funds Sorted by MAX(N)K

Each month t from 1998/9 to 2010/12, funds are sorted into deciles based on the average of N (N=1, 2, 3, 4, 5) highest daily style-adjusted returns (MAX $(N)_K$) over the previous K months where K=3,6,12 months. These decile portfolios are held for a month after portfolio formation. This table reports equally weighted monthly portfolio flows for each decile in month t+1.Decile 1(10) is the group of funds with the lowest (highest) maximum multi-day returns. Flow is the growth rate of assets under management due to new investments The last row (HM-LM) presents the difference of flow ratios between the High MAX (decile 10) and Low MAX (decile 1) groups. Newey and West (1987) adjusted t-statistics (up to 6 lags) are reported in parenthesis. All values except the t-statistics are given in percentage terms. Panel A reports flows for decile portfolios calculated based on MAX measured over the previous 3 months. Panel B reports flows for decile portfolios sorted on MAX measured over the previous 6 months, and Panel C reports flows for decile portfolio sorted on MAX measured over the previous 12 months. Panel D presents flows for decile portfolios sorted on the average of single day MAX over the previous 2 ,3, 6, 12, and 18 months.

Panel A: MAX calculated over the previous 3 months

Decile	$Max[1]_3$	$Max[2]_3$	$Max[3]_3$	$Max(4)_3$	$Max(5)_3$
Low MAX	0.32	0.28	0.29	0.28	0.27
2	0.35	0.36	0.33	0.33	0.32
3	0.44	0.43	0.42	0.42	0.42
4	0.54	0.56	0.53	0.51	0.51
5	0.51	0.49	0.53	0.54	0.54
6	0.58	0.56	0.54	0.53	0.52
7	0.62	0.60	0.60	0.61	0.61
8	0.70	0.65	0.68	0.67	0.67
9	0.78	0.80	0.82	0.83	0.84
High MAX	1.05	1.13	1.14	1.17	1.18
HM-LM	0.73	0.85	0.85	0.89	0.91
	(4.25)	***(4.63)***	***(4.38)***	***(4.59)***	***(4.65)***

Panel B: MAX calculated over the previous 6 months

Decile	$Max[1]_6$	$Max[2]_6$	$Max[3]_6$	$Max[4]_6$	$Max[5]_6$
Low MAX	0.30	0.28	0.28	0.27	0.27
2	0.36	0.34	0.32	0.30	0.29
3	0.46	0.47	0.46	0.47	0.46
4	0.47	0.51	0.52	0.53	0.53
5	0.60	0.56	0.54	0.51	0.50
6	0.60	0.57	0.59	0.58	0.57
7	0.59	0.59	0.59	0.60	0.59
8	0.67	0.66	0.66	0.70	0.72
9	0.86	0.88	0.89	0.87	0.85
High MAX	1.01	1.06	1.08	1.10	1.13
HM-LM	0.72	0.78	0.80	0.82	0.86
	(4.26)	***(4.16)***	***(4.07)***	***(5.45)***	***(5.54)***

Panel C: MAX calculated over the previous 12 months

Decile	$Max[1]_{12}$	$Max[2]_{12}$	$Max[3]_{12}$	$Max[4]_{12}$	$Max[5]_{12}$
Low MAX	0.27	0.25	0.25	0.24	0.24
2	0.31	0.31	0.28	0.26	0.25
3	0.48	0.47	0.48	0.48	0.48
4	0.51	0.51	0.50	0.52	0.54
5	0.58	0.55	0.56	0.56	0.54
6	0.59	0.61	0.58	0.55	0.55
7	0.65	0.62	0.61	0.62	0.62
8	0.63	0.62	0.64	0.63	0.66
9	0.77	0.85	0.86	0.86	0.88
High MAX	0.93	0.92	0.94	0.97	0.95
HM-LM	0.66	0.67	0.69	0.73	0.72
	(5.11)	***(4.78)***	***(4.95)***	***(5.08)***	***(4.96)***

Panel D: Average of MAX over the previous 2, 3, 6, 12, 18 months

Decile	$AMax_2$	$AMax_3$	$AMax_6$	$AMax_{12}$	$AMax_{18}$
Low MAX	0.28	0.27	0.26	0.21	0.13
2	0.37	0.36	0.30	0.25	0.29
3	0.41	0.41	0.42	0.42	0.46
4	0.49	0.47	0.50	0.51	0.47
5	0.56	0.53	0.51	0.50	0.42
6	0.55	0.57	0.58	0.59	0.59
7	0.59	0.56	0.61	0.64	0.68
8	0.66	0.70	0.67	0.66	0.69
9	0.83	0.82	0.84	0.87	0.81
High MAX	1.16	1.18	1.22	1.18	1.15
HM-LM	0.87	0.92	0.96	0.97	1.02
	(4.68)	***(4.74)***	***(4.54)***	***(4.24)***	***(4.67)***

suggests that investors are increasingly attracted to funds if they have high MAX on average over past periods. Overall, using different horizons and the number of days to proxy extreme returns produce similar results. In the rest of the paper I focus on MAX- single day maximum return over a month- , but the results are very similar if I use one of these alternative definitions.

In univariate sorts, I show that there is a positive relationship between flows and MAX. However, as presented in the previous section, fund characteristics such as style-adjusted returns, size, expenses, turnover, volatility, and skewness vary across MAX portfolios. These variables might also influence flows. To determine whether the positive relation between flows and MAX holds after controlling these fund characteristics, I conduct bivariate (sequential) sorts. These sorts also give insights about how the flow-MAX relation changes across different types of funds. Specifically, in each month, I first rank all funds in ascending order according to different fund characteristics and allocate them into three groups- Bottom, Mid, Top-based upon a 20:60:20 split. Then within each group, I further sort the funds into deciles based on MAX within the same month. This creates 30 portfolios, which are held for a month after portfolio formation. I alternatively form 25 portfolios by 5x5 sequential (dependent) sorts as well as 100 portfolios by 10x10 sequential sorts. The results are similar and do not change any inference. Moreover, instead of sequential sorts, I also construct portfolios using independent sorts. The results are again similar to those in sequential sorts and available upon request. Table 1.5 reports average flows in the subsequent month for the 30 portfolios based on the 20:60:20 approach along with the flow difference between high MAX (decile 10) and low MAX (decile 1) portfolios.

The results in Table 1.5 suggest that the flow-MAX relation is present in all return controlled groups. However, the effect is much stronger for winners than for losers. Among losers, the flow difference between high MAX (decile 10) and low MAX (decile 1) portfolio is 0.37% per month (t-statistics: 2.04), which is half of the unconditional effect of MAX (0.84%) reported in Table 1.3. Among winners, the flow difference is much greater with a value of 0.80% per month and is highly statistically significant. Table 1.5 documents that high MAX funds are more likely to have higher returns in the same month (1.55% vs.-0.03%). If investors form expectations about future high payoff states based on previous MAX, then the strong effect among winners makes sense since high MAX funds are dominated by winners. The results also suggest that the flow-MAX relation is stronger among relatively bigger funds, funds with higher expense ratios, turnover, and volatility, younger funds, and funds with lower skewness.

<u>*Multivariate Regressions*</u>

In this section, I adopt a regression approach based on Fama and MacBeth (1973), which controls for multiple fund characteristics simultaneously. Specifically, I run the following regression:

$$\mathrm{Flow}_{i,t+1} = \alpha + \beta_{i,t} \times \mathrm{MAX}_{i,t} + \delta_{i,t} \times \mathrm{X}_{i,t} + \varepsilon_{i,t} \qquad (1.2)$$

where $\mathrm{Flow}_{i,t+1}$ is the flow ratio for fund i in month t+1 and $\mathrm{X}_{i,t}$ is a vector of control variables for fund i in month t. The control variables are the logarithm of fund size, the logarithm of fund age, Non12b-1 fees, 12b-1 fees, turnover, front load, rear load, fund flows, and past performance. Among others, Chevalier and Ellison (1997), Sirri and Tufano (1998), Del Guercio and Tkac (2008), and Barber, Odean and Zheng (2005) show that these variables influence future fund flows.

Table 1.5. Average Monthly Flow of Funds Sorted on MAX and Fund Characteristics

Each month t from 1998/9 to 2010/12, funds are ranked in ascending order according to different fund characteristics in month t and are allocated into three groups. Bottom, Mid, Top-based upon a 20:60:20 split. Bottom (Top) refers to the group of funds that are in the bottom (top) 20% of all funds ranked according to the fund characteristic. Mid includes funds that are in the 2nd-4th quintiles according to the fund characteristics. Then within each group, I further sort the funds into deciles based on MAX within the same month. This creates 30 portfolios which are held for a month after portfolio formation. This table reports equally weighted monthly portfolio flows for each decile in month t+1. Decile 1(10) is the group of funds with the lowest (highest) MAX. Flow is the growth rate of assets under management due to new investments.. The last row (HM-LM) presents the difference of flow ratios between the High Max (decile 10) and Low Max (decile 1) groups. Newey and West (1987) adjusted t-statistics (up to 6 lags) are reported in parenthesis. All values except the t-statistics are given in percentage terms.

	Return			Size			Age			Turnover		
	Bottom	Mid	Top	Bottom	Mid	Top	Bottom	Mid	Top	Bottom	Mid	Top
Low MAX	0.16	0.34	0.87	1.30	0.13	-0.14	1.58	0.19	-0.77	0.45	0.14	0.74
2	0.11	0.39	1.10	1.52	0.21	-0.14	1.82	0.33	-0.75	0.48	0.26	0.73
3	0.27	0.46	1.00	1.40	0.25	-0.06	1.86	0.32	-0.70	0.51	0.33	0.73
4	0.20	0.43	1.07	1.56	0.29	0.04	1.93	0.45	-0.69	0.56	0.40	0.74
5	0.24	0.50	1.11	1.65	0.37	0.08	1.99	0.43	-0.55	0.60	0.45	0.67
6	0.29	0.55	1.10	1.65	0.37	0.15	2.00	0.48	-0.59	0.68	0.49	0.80
7	0.23	0.51	1.35	1.52	0.44	0.21	2.09	0.52	-0.50	0.79	0.50	0.90
8	0.15	0.58	1.43	1.66	0.47	0.23	2.05	0.54	-0.50	0.77	0.52	0.84
9	0.24	0.54	1.49	1.58	0.64	0.33	2.34	0.65	-0.33	0.93	0.64	0.97
High MAX	0.53	0.69	1.66	1.75	1.02	0.65	2.80	0.91	-0.03	0.93	0.84	1.84
	0.37	0.35	0.80	0.45	0.88	0.79	1.22	0.72	0.74	0.47	0.70	1.11
	(2.04)	***(2.86)***	***(4.99)***	***(2.16)***	***(4.72)***	***(6.48)***	***(5.28)***	***(5.15)***	***(7.81)***	***(3.01)***	***(4.17)***	***(3.57)***

	Expense Ratio			12b-1			Volatility			Skewness		
	Bottom	Mid	Top	Bottom	Mid	Top	Bottom	Mid	Top	Bottom	Mid	Top
Low MAX	0.55	0.23	-0.10	0.65	0.38	-0.35	0.55	0.25	0.31	0.81	0.30	0.06
2	0.61	0.39	-0.12	0.76	0.48	-0.20	0.53	0.30	0.40	0.96	0.36	0.09
3	0.73	0.47	-0.14	0.65	0.54	-0.16	0.58	0.43	0.44	0.98	0.36	0.01
4	0.67	0.60	0.00	0.75	0.67	-0.13	0.74	0.40	0.48	1.05	0.42	0.05
5	0.73	0.65	0.02	0.75	0.66	0.00	0.67	0.50	0.62	1.01	0.51	0.16
6	0.82	0.70	0.04	0.76	0.71	0.07	0.90	0.50	0.64	1.07	0.55	0.14
7	0.74	0.78	0.10	0.86	0.76	0.05	0.79	0.52	0.73	1.22	0.55	0.21
8	0.81	0.76	0.18	0.79	0.79	0.05	0.89	0.61	0.78	1.35	0.60	0.25
9	0.80	0.99	0.39	0.86	0.97	0.12	0.87	0.67	0.98	1.45	0.64	0.40
High MAX	1.06	1.27	0.82	1.06	1.32	0.50	1.04	0.84	1.52	1.74	0.96	0.63
	0.51	1.04	0.92	0.41	0.94	0.84	0.49	0.59	1.21	0.93	0.66	0.57
	(3.42)	***(6.75)***	***(3.84)***	***(2.63)***	***(5.90)***	***(3.43)***	***(4.55)***	***(4.91)***	***(4.77)***	***(4.73)***	***(4.69)***	***(2.80)***

MAX may increase the visibility of a fund and this visibility may attract investors' attention. Even though I do not completely rule out this explanation, I use three additional control variables including family size, large family dummy, and star fund dummy, which are arguably better proxies for visibility in the mutual fund industry.[18] Most funds belong to families.[19]Because of brand-building and marketing benefits; belonging to a large family increases the promotion and visibility of a fund, which may attract investors. Therefore, I use log of family size and a dummy variable that equals to one if a fund is a member of a large family.[20] Moreover, funds with stellar performance receive media attention and attract investors (Nanda, Vikram and Zheng (2004), Del Guercio and Tkac (2008)). To control for this "star" effect, I rank funds each month according Carhart's (1997) alphas over the previous 60 months. Funds in the top 5% are defined as star funds.[21] All specifications include style-dummies (based on 12 Lipper classifications) to control for the style fixed effects and to allow for changing investor tastes for funds with different investment styles (Bergstresser and Poterba (2002)).

While MAX is arguably a theoretically motivated measure, there is still a concern that it might proxy for volatility. In particular, volatile funds are more likely to exhibit extreme returns as suggested by the high positive correlation between two variables. Equally important is to examine the ability of MAX to explain flows after controlling for

[18] Fund size and 12b-1 expenses (i.e. marketing, advertising, and brokerage costs) are other control variables, which are related to visibility of a fund. Sirri and Tufano (1998) find that media coverage, hence visibility, increases with fund size and volatility of returns.

[19] In my sample, 88% of funds are affiliated with a fund complex.

[20] For each month, a large family is defined as the family with total net assets in the top 10^{th} percentile. Other cut-off points such as the top 5^{th} or 1^{th} percentile produce similar results.

[21] Using Carhart's (1997) alphas over the previous 60, 36 and 24 produce similar results. I also rank funds within each style based on returns over the previous 12 months. Results are again similar.

skewness. Thus, I use volatility and skewness to control whether MAX is a proxy for these variables.[22] I adjust Fama and MacBeth's (1973) t-statistics for heteroskedasticity and autocorrelation up to 6 lags.[23]

The extant literature (Chevalier and Ellison (1997), Sirri and Tufano (1998)) documents convexity in the relation between fund flows and past performance. To capture this feature, I adopt two different methods. The first one is outlined in Sirri and Tufano (1998). Specifically, each month I rank all funds according to their one month returns within each investment style and assign them fractional ranks uniformly distributed between 0 (worst performance) and 1 (best performance). These ranks represent a fund's percentile performance relative to other funds in that month. I then partition relative performance into three groups:

$$\begin{aligned} \text{Low}_{i,t} &= \text{Min}(\text{Rank}_{i,t}, 0.2) \\ \text{Mid}_{i,t} &= \text{Min}(\text{Rank}_{i,t} - \text{Low}_{i,t}, 0.6) \\ \text{High}_{i,t} &= \text{Rank}_{i,t} - \text{Mid}_{i,t} - \text{Low}_{i,t} \end{aligned} \tag{1.3}$$

[22] In Table 1.6, volatility is calculated using daily returns over a month, while daily returns over the previous 12 months are used for skewness. However, the results are similar if volatility is measured using daily returns over the previous 3, 6 or 12 months, or if skewness is measured over the previous 1, 3 or 6 months. Moreover, in section V, I investigate the relation between MAX and idiosyncratic volatility and idiosyncratic skewness. The effect of MAX on flows is robust to control for idiosyncratic volatility and skewness.

[23] To validate the selection of 6 lags, I use the Pontiff (1996) method to test the Fama–MacBeth standard errors for potential higher-order serial correlation. Specifically, I estimate an autoregressive model using the time series of MAX coefficient estimates. The order of the autoregressive model is chosen such that its Durbin-Watson (DW) statistic is close to two. The DW statistics for autoregressive models of AR(1) to AR(5) range from 1.96 to 1.98 suggesting a slight positive autocorrelation. I find that six lags are sufficient to eliminate the serial correlation in errors (DW = 2.0006). Pontiff (1996) use the standard errors of the intercept in autoregressive models as the autocorrelation corrected standard errors of the coefficient estimates. Using Newey and West’s (1987) correction of 6 lags, I find more conservative standard errors. Moreover, although I correct standard errors for potential heteroskedasticity and autocorrelation, I also use a panel regression with standard errors clustered by both month and fund (Peterson (2009)). Results are in Table 1.17 , and are qualitatively similar to the results obtained from Fama-Macbeth regressions.

These three groups are added into the regression as control variables. In the second method, I directly use style-adjusted returns and the squared value of these returns in the regressions to capture the nonlinearities in the flow-performance relation. The first specification relies on the normalized performance, the second one uses absolute performance.

The results are presented in Table 1.6.[24] When the normalized performance is used (column-I), the coefficient of MAX is 0.146 and is statistically significant at the 1% level. According to Table 1.3 , the average of MAX for funds in the top decile is 1.84%, while it is 0.23% for funds in the bottom decile. Multiplying the spread between these two deciles by the coefficient yields an estimate of additional flows of 2.82% per year, which is 13.2% of the average annual flows (21.3% per year). [25] Controlling for both volatility and skewness (column-VII), the effect of MAX further increases to 0.239, which translates an incremental gain of 4.62% in flows per year, which is 21.7% of the average annual flows. [26] When the absolute performance is used, the coefficients of MAX range between 0.110 and 0.195, and are all highly significant. [27] On the other hand, volatility loads negatively and has a significant coefficient, which is consistent with Huang, Wei, and Yang (2007). Similarly, skewness is also negatively related to future

[24] I multiply the coefficient of size, family size, and turnover by 100.

[25] To put it into context, moving 5^{th} percentile in the highest performance category (for example from the 85^{th} to the 90^{th} percentile) increases the flow by 2.1% per year.

[26] Instead of partitioning the fractional ranks into three groups- Low, Mid, High- I divide them into ten deciles perform the same tests by using 10 deciles in order to assess the robustness of the results to high percentiles (i.e. funds in the top 10%). The coefficient of MAX is 0.136 with a standard error of 0.026. When both volatility and skewness are included into the regressions, the coefficient of MAX increases to 0.228 with a standard error of 0.037.

[27] I also use an interaction term of volatility and skewness, with the idea that not skewness per se but skewness with high volatility may better reflect lottery-like preferences (Kumar (2009)). The coefficient of MAX remains highly significant, while the interaction term is not significant.

Table 1.6. Flow-MAX Relation

This table reports the parameter estimates from Fama and MacBeth (1973) regressions for funds from 1998/10 to 2010/12. The dependent variable is the flow ratio. The independent variables are MAX over the previous month and a set of control variables including volatility and skewness of fund returns, flows, the logarithm of size, the logarithm of age, the logarithm of family size, 12b-1 fees, Non12b-1 fees (as measured by the difference between expense ratios and 12b-1 fees), rear loads, front load, turnover, star-fund dummy, big-family dummy, and fund performance. All control variables are lagged at least one month and all specifications include style-dummies. Standard errors (in parenthesis) are corrected for heteroskedasticity and autocorrelation (up to 6 lags) using Newey and West (1987).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MAX	0.146***	0.110***	0.230***	0.187***	0.156***	0.119***	0.239***	0.195***
	(0.027)	(0.030)	(0.041)	(0.042)	(0.026)	(0.028)	(0.039)	(0.040)
Volatility			-0.573***	-0.606***			-0.586***	-0.627***
			(0.167)	(0.170)			(0.158)	(0.160)
Skewness					-0.004***	-0.004***	-0.003***	-0.003***
					(0.001)	(0.001)	(0.001)	(0.001)
Flow	0.347***	0.347***	0.345***	0.345***	0.346***	0.346***	0.345***	0.345***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Log(Size)	-0.109***	-0.109***	-0.111***	-0.111***	-0.109***	-0.109***	-0.111***	-0.111***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Log(Fsize)	0.041***	0.040***	0.045***	0.045***	0.042***	0.042***	0.047***	0.046***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Big Family	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Star Fund	0.011***	0.011***	0.011***	0.011***	0.011***	0.011***	0.011***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log(Age)	-0.580***	-0.579***	-0.580***	-0.579***	-0.585***	-0.584***	-0.585***	-0.584***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.030)	(0.030)	(0.030)	(0.030)
Non12b-1	-0.413***	-0.409***	-0.407***	-0.400***	-0.415***	-0.409***	-0.408***	-0.400***
	(0.045)	(0.045)	(0.045)	(0.044)	(0.046)	(0.045)	(0.046)	(0.045)
12b-1	-0.705***	-0.708***	-0.706***	-0.708***	-0.732***	-0.734***	-0.733***	-0.735***
	(0.104)	(0.103)	(0.106)	(0.106)	(0.101)	(0.101)	(0.104)	(0.103)
Rear Load	-0.135***	-0.135***	-0.136***	-0.136***	-0.140***	-0.140***	-0.141***	-0.141***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.016)	(0.015)
Front Load	-0.023***	-0.023***	-0.023***	-0.023***	-0.025***	-0.025***	-0.025***	-0.025***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Turnover	-0.030	-0.035*	-0.024	-0.028	-0.035*	-0.040**	-0.026	-0.031*
	(0.020)	(0.020)	(0.018)	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)
Low	0.012***		0.015***		0.012***		0.015***	
	(0.002)		(0.002)		(0.002)		(0.002)	
Mid	0.004***		0.005***		0.005***		0.005***	
	(0.001)		(0.001)		(0.001)		(0.001)	
High	0.032***		0.035***		0.032***		0.035***	
	(0.002)		(0.002)		(0.002)		(0.002)	
Perf		0.167***		0.188***		0.170***		0.190***
		(0.017)		(0.016)		(0.017)		(0.015)
(Perf)2		1.369***		1.437***		1.376***		1.475***
		(0.301)		(0.323)		(0.295)		(0.319)
N	626623	626623	626620	626620	618708	618708	618705	618705
R-sq	0.21	0.21	0.21	0.21	0.21	0.21	0.21	0.21

fund flows. Overall, these results suggest that the effect of MAX is not only robust to controls for volatility and skewness, but is also the opposite of volatility and skewness effects.

In general, the coefficients on the control variables are consistent with previous research. For instance, as documented in Sirri and Tufano (1998), the flow-performance relation is convex no matter whether the performance is normalized or not. Size and age negatively affect performance (Chevalier and Ellison (1997), Huang, Wei, and Yan (2007)). Consistent with Nanda, Wang, and Zheng (2004), star performance results in greater inflows. Family size and belonging to a large family positively affect flows, presumably due to a branding and visibility effect. For the period 1993-1999, Barber, Odean, and Zheng (2005) show a positive relation between 12b-1 fees and flows, and interpret this evidence as the role of advertising in garnering new money from investors. Instead, I document a negative relation 12b-1 fees and flows as in Gil-Bazo and Ruiz-Verdύ (2009). One possible explanation for the discrepancy in those results is that the increasingly easy access to financial information and the decreasing participation and search costs over time might mitigate the effect of advertising and distribution services on flows.

If investors prefer high MAX funds due to their lottery like preferences and immediately invest in these funds after observing it, then the timing of the MAX in a given month might be important in determining the flow amount next month. For instance, funds which experience MAX over the early days of a month would have part of their inflows over the same month MAX measured, whereas funds which experience MAX in the later days of a month experience majority of the flows over the next month.

Since I measure flow using monthly data, then months which experience MAX in the later days of a month should have more inflows. To examine this possibility, I create a dummy variable (MAXTIME), which is equal to 1 if MAX occurs in the second half of the month (i.e. week 3 or 4) or 0 otherwise. I included the interaction term of MAX and MAXTIME in Fama and Macbeth's (1973) regressions will full set of control variables. The interaction term of MAX and MAXTIME is significant with a value of 0.114, while the coefficient of MAX drops to 0.099 but remains significant. This suggests that when MAX occurs in the later days of a month, fund experience more inflow next month as majority of the flows will be captured by the next month flow, which is in line with the arguments above. [28]

There is considerable skewness in fund size. As reported in Table 1.1, average total net assets (TNA) of $490.42 million, while the corresponding median value is 52.79 million. Clifford and Jordan (2011) argue that due to the skewness in fund size, what is true for an average fund may not true for an average fund investor. For instance, Gruber (1996) shows that investors buy funds that subsequently outperform, while selling funds that subsequently underperform. Hence, investors appear to be "smart" in their fund choices. [29] Clifford and Jordan (2011) document that the smart money effect is confined to only small funds (i.e. funds with size below median of the distribution) that have limited economic importance, and does not exist in big funds, which an average fund

---

[28] Carhart, Kaniel, Musto, and Reed (2002) find that fund managers aggressively buy the stocks they already have at the end of the quarter to push up stock prices in order to have high NAVs at period ends and increase their flows. MAX in returns may be produced through this process as managers' actions will push the stock prices up. However, the effect of MAX in my sample does not change when I eliminate the quarter end months from my sample.

[29] Even though Sapp and Tiwari (2004) argue that this effect is due to momentum, Keswani and Stolin (2008) find evidence of "smartness" when studying the flows of both US and UK investors.

investor holds.[30] Hence, a result that is significant for funds collectively may lack economic significance for fund investors unless the effect is observed in bigger funds as well. To investigate whether the effect of MAX is solely driven by small funds, for each month I create a dummy variable (Medsize), which takes the value of 1 if the fund has total net assets greater than the median size of all funds, or 0 otherwise. I use the interaction of this variable with MAX and dummy variable in my regressions. Other control variables are same as in Table 1.6. Panel A of Table 1.7 presents the results.

For brevity, I omit the coefficients of control variables except volatility and skewness. The interaction terms in all specifications are positive, albeit not statistically significant, ranging from 0.008 to 0.028. Next, I estimate two separate regressions (i.e. within above-median and below-median groups), thereby allowing all the coefficients to vary. In all specifications (Panel B of Table 1.7), the coefficients of MAX in the above-median group are highly statistically significant and greater than the ones in the below-median group. The coefficients of MAX in the below-median group are also significant in all but one specification. [31] The standard deviation of MAX is 1.16% in the above-median groups, while it is 0.7% in the below-median group. This suggests that the economic impact of MAX is actually greater than what is reported in Table 1.6. Overall,

[30] The authors also investigate the convexity in flow-performance relation and the negative relation between fund performance and expense ratio documented by Gil-Bazo and Ruiz-Verdύ (2009). As in smart money effect, the convexity in flow-performance relation and the relation between fund performance and expense ratio

[31] One point of concern is that the survivor bias still exists in the CRSP data (Elton, Gruber, and Blake (2001)). While this bias particularly affects the absolute level of fund performance but it is not likely to bias the estimated coefficients that measure the sensitivity of flows to MAX. If MAX funds pursue risky strategies and die going forward, this would probably make it harder to find a negative relation between performance and MAX. Moreover, my results show that the effect of MAX is not confined to small funds which are more prone to pursue risky strategies. Contrary, the effect of MAX is more prevalent in bigger funds where **average** investors are concentrated in (Clifford and Jordan (2011)) and which are less likely to die.

Table 1.7. Flow-MAX Relation within Different Size Groups

This table reports the parameter estimates from Fama and MacBeth (1973) regressions of monthly observations for funds from 1998/10 to 2010/12. The dependent variable is the flow ratio. In Panel A, Medsize is a dummy variable that takes the value of 1 if the fund has total net assets greater than the median size of all funds, zero otherwise. Panel B presents the results of separate regression for the above and below median groups. Other control variables (omitted) are same as in Table 1.6. Standard errors (in parenthesis) are corrected for heteroskedasticity and autocorrelation (up to 6 lags) using Newey and West (1987).

Panel A: Full sample regressions with the interaction term

	Piecewise Linear Regression				Quadratic Regression			
MAX	0.139***	0.227***	0.147***	0.236***	0.098**	0.180***	0.108***	0.187***
	(0.038)	(0.042)	(0.038)	(0.041)	(0.038)	(0.042)	(0.037)	(0.040)
MAX*Medsize	0.016	0.008	0.021	0.014	0.027	0.019	0.028	0.022
	(0.064)	(0.066)	(0.064)	(0.066)	(0.062)	(0.064)	(0.062)	(0.064)
Volatility		-0.578***		-0.592***		-0.614***		-0.635***
		(0.165)		(0.156)		(0.168)		(0.158)
			-0.004***	-0.003***			-0.004***	-0.003***
Skewness			(0.001)	(0.001)			(0.001)	(0.001)

Panel B: Separate regressions within above-median and below-median groups

	Above Median							
	Piecewise Linear Regression				Quadratic Regression			
MAX	0.161***	0.227***	0.176***	0.241***	0.119***	0.178***	0.133***	0.191***
	(0.044)	(0.056)	(0.043)	(0.053)	(0.042)	(0.052)	(0.041)	(0.050)
Volatility		-0.500***		-0.513***		-0.506***		-0.520***
		(0.143)		(0.132)		(0.146)		(0.135)
Skewness			-0.005***	-0.004***			-0.005***	-0.004***
			(0.001)	(0.001)			(0.001)	(0.001)

	Below Median							
	Piecewise Linear Regression				Quadratic Regression			
MAX	0.097**	0.180***	0.107**	0.188***	0.062	0.141***	0.070*	0.147***
	(0.044)	(0.048)	(0.043)	(0.047)	(0.041)	(0.050)	(0.039)	(0.047)
Volaility		-0.525***		-0.549***		-0.604***		-0.637***
		(0.187)		(0.182)		(0.189)		(0.182)
Skewness			-0.004***	-0.003**			-0.004***	-0.003**
			(0.001)	(0.001)			(0.001)	(0.001)

my analysis suggests that the effect of MAX is not confined to small funds. Indeed, if anything, it is stronger in bigger funds.

The findings of Bali, Cakici, and Whitelaw (2011) might raise the concern that fund investors are attracted to underlying stocks' MAX rather than the funds' MAX itself. I argue that this is a remote possibility. First, investors do not observe the underlying individual stocks of mutual funds on daily or monthly basis like fund returns. Second, fund returns are not exactly driven only by common stocks in the fund portfolio. These returns are net of expenses and reflect the valuation of overall fund portfolio which may include bonds, cash, and other instruments. Third, investors may be attracted to funds based on stocks' MAX if stock attributes relating to MAX automatically translate into the same portfolio attributes when these stocks are put into a fund.

However, the existing literature already documents that individual stock characteristics may not translate into portfolio characteristics. For example while the aggregate market returns are negatively skewed, individual stock returns exhibit positive skewness (Hong and Stein (2003), Duffee (1995)). The disconnection between skewness of individual stocks and portfolio skewness is due to the comovement in one stock's return with the remaining stocks in the portfolio. As shown in Albuquerque (2012), in a portfolio, the comovement term dominates the weighted average of individual stock skewness as the number of stocks in the portfolio increases.[32] Hence, it is likely to form a portfolio that lacks a certain characteristics, even though the underlying assets have the

[32] The literature suggests many reasons why this term may be negative, which subsequently leads to a portfolio with negatively skewed returns while the stocks in the portfolio are positively skewed. Albuquerque (2012) develops a model in which cross-sectional heterogeneity in firm announcement events can lead to conditional asymmetric stock return correlations and negative skewness in aggregate returns. Hong and Stein (2003) argue that one likely factor is the tendency for managers to release negative firm-specific information in a gradual piecemeal fashion.

very same characteristics. My main variable, MAX, uses only one day return of a portfolio over an extended period of time (i.e. one month). Even though the portfolio may consist of some stocks that produce MAX over a month, the effect of individual stock MAX in the portfolio may not appear in portfolio returns due to correlation of that stock with the remaining stocks in that portfolio.[33] Therefore, given that MAX of an individual stock may be weakened in a portfolio; it is less likely that an investor cares about individual stock MAX in a fund.

Retail vs. Non-Retail Funds

So far I have documented a positive and significant relation between MAX and future flows, which is robust to alternative definitions of MAX and controls for other fund characteristics. Moreover, this relation is stronger in bigger funds. As argued before, one plausible explanation for my findings is that some investors are attracted to lottery like-payoffs as modeled in Barberis and Huang (2008) and Brunnermeier, Gollier, and Parker (2007). If flow-MAX relation can be attributed to this phenomena, then I expect to find a stronger relation between MAX and future flows among retail funds as retail investors are documented to more likely exhibit this kind of preference (Kumar (2009), Åstebro, Mata, and Santos-Pinto (2009)). To examine this idea, I divide all funds into

[33] For an illustration, consider the following example with three stocks:

Period 1	Period2
A: 15% B:-15% C: 15%	A:-15% B: 15% C: 0%

For all of the three stocks above (A, B, C), MAX measured over two periods is 15%. Yet, if you invest 50% on A and B, the portfolio will have 0% in both periods, while if you invest 50% in A and C, then the portfolio will have a return of 15% in period 1 and 7.5% in period 2. Hence, Max for portfolio 1 will be 0%, while Max for portfolio 2_will be 15%. For illustrative purposes, I use two perfectly negatively correlated stocks; however, the same argument may apply to any number of stocks with negative cross-correlations.

retail and non-retail funds using the "*retail_fund*" flag in CRSP.[34] This flag is available from 1999.[35]

Table 1.8 presents Fama and MacBeth's (1973) regression estimates within retail and non-retail funds. Panel A reports the coefficients from piecewise linear regressions and Panel B shows the coefficients from quadratic regressions. Within retail funds, the coefficient of MAX is 0.192 when normalized performance is used and 0.141 when absolute performance is used. Both values are greater than the estimates in Table 1.6 and highly statistically significant at 1% level. On the other hand, the corresponding coefficients for MAX in non-retail funds are 0.079 and 0.086. Strikingly, none of these estimates are statistically significant at any conventional level. Within retail funds, the coefficients of MAX increase to 0.230 in piecewise regressions and 0.179 in quadratic regressions after controlling for volatility and skewness. They are all statistically significant and greater than the coefficients in the regressions that do not use volatility or skewness as control variables. Again, in non-retail funds, there is no evidence of a significant relation between MAX and fund flows when volatility and skewness are included. One exception is that MAX becomes significant in non-retail funds after adding both volatility and skewness into piecewise regressions. But this result is not robust to using quadratic specification. Therefore, my analysis suggests that the flow-MAX relation is more pronounced among retail funds. This is consistent with my priori that the

[34] There are, on average, 3177 retail funds and 1766 non-retail funds each month in my sample.

[35] According to CRSP, retail fund flag determines whether the fund is a retail fund, which caters primarily to individual investors. One drawback of this classification is that it may not be a precise identifier of investor type. For instance, investment decisions in a 401K plan are eventually taken by individual investors. However, capital flows may combine flows of either an institutional or retail fund. Nevertheless, it seems reasonable to argue that classification of funds given by CRSP implies differences in investor composition of the two types of fund given the fact that majority of retail fund investors are regular individuals. Moreover, institutional investors can be expected to invest more in institutional funds to exploit the benefits of institutional funds such as low expenses.

Table 1.8. Flow-MAX Relation within Retail vs. Non-Retail Funds

This table reports the parameter estimates from Fama and MacBeth (1973) regressions from 1998/10 to 2010/12 for retail and non-retail funds. The dependent variable is the flow ratio. The independent variables are MAX over the previous month. Control variables are flows, the logarithm of size, the logarithm of age, 12b-1 fees, Non12b-1 fees, rear loads, front load, turnover and fund performance. All control variables are lagged at least one month. All specifications include style-dummies. Panel A shows the estimates from piecewise linear regressions, while Panel B reports the estimates from quadratic regressions. Newey and West's (1987) standard errors (6 lags), are in parenthesis.

Panel A: Piecewise Linear Regressions

	Retail Funds				Non-Retail Funds			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MAX	0.192***	0.241***	0.175***	0.230***	0.079	0.146	0.093*	0.218**
	(0.037)	(0.044)	(0.030)	(0.041)	(0.051)	(0.091)	(0.051)	(0.101)
Volatility		-0.448**		-0.523***		-0.587***		-0.635***
		(0.183)		(0.154)		(0.208)		(0.207)
Skewness			-0.004***	-0.003**			-0.007***	-0.006***
			(0.001)	(0.001)			(0.002)	(0.002)
Flow	0.374***	0.370***	0.382***	0.380***	0.282***	0.281***	0.279***	0.278***
	(0.018)	(0.018)	(0.017)	(0.017)	(0.009)	(0.009)	(0.009)	(0.009)
Log(Size)	-0.150***	-0.176***	-0.117***	-0.121***	-0.060**	-0.060**	-0.059**	-0.059**
	(0.036)	(0.053)	(0.017)	(0.017)	(0.027)	(0.027)	(0.028)	(0.028)
Log(Age)	-0.513***	-0.473***	-0.534***	-0.531***	-0.701***	-0.706***	-0.712***	-0.717***
	(0.037)	(0.064)	(0.030)	(0.030)	(0.041)	(0.041)	(0.042)	(0.042)
Log(Fsize)	0.040**	0.050**	0.024***	0.028***	0.114***	0.118***	0.117***	0.121***
	(0.018)	(0.022)	(0.009)	(0.009)	(0.012)	(0.012)	(0.012)	(0.012)
Big Family	0.002***	0.002***	0.002***	0.002***	0.002**	0.001*	0.002*	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Star Fund	0.011***	0.011***	0.011***	0.011***	0.009***	0.009***	0.009***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Non12b-1	-0.448***	-0.444***	-0.445***	-0.437***	-0.313***	-0.301***	-0.315***	-0.302***
	(0.053)	(0.055)	(0.049)	(0.048)	(0.064)	(0.063)	(0.066)	(0.066)
12b-1	-0.632***	-0.681***	-0.603***	-0.606***	0.098	0.091	0.091	0.084
	(0.098)	(0.117)	(0.096)	(0.099)	(0.203)	(0.202)	(0.207)	(0.207)
Rear Load	-0.095***	-0.081*	-0.124***	-0.125***	-0.103***	-0.103***	-0.106***	-0.107***
	(0.033)	(0.046)	(0.014)	(0.015)	(0.036)	(0.036)	(0.037)	(0.037)
Front Load	-0.017	-0.002	-0.005	-0.006	-0.083***	-0.079**	-0.086***	-0.082***
	(0.012)	(0.010)	(0.005)	(0.005)	(0.031)	(0.031)	(0.031)	(0.031)
Turnover	-0.097	-0.079	-0.020	-0.012	-0.070***	-0.061**	-0.081***	-0.070***
	(0.076)	(0.068)	(0.020)	(0.019)	(0.026)	(0.024)	(0.025)	(0.024)
Low	0.012***	0.011*	0.014***	0.017***	0.006*	0.008**	0.006*	0.008**
	(0.004)	(0.006)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)
Mid	0.004***	0.006***	0.005***	0.005***	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
High	0.038***	0.038***	0.034***	0.037***	0.023***	0.024***	0.023***	0.024***
	(0.004)	(0.003)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)	(0.004)
N	435077	435074	434921	434918	191546	191546	183787	183787
R-sq	0.24	0.25	0.25	0.25	0.14	0.14	0.14	0.14

Table 1.8 (Continued)

Panel B: Quadratic Regressions

	Retail Funds				Non-Retail Funds			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	8)
MAX	0.141***	0.183***	0.133***	0.179***	0.086	0.076	0.096*	0.152
	(0.036)	(0.044)	(0.031)	(0.041)	(0.053)	(0.094)	(0.053)	(0.105)
Volatility		-0.528***		-0.565***		-0.624***		-0.671***
		(0.174)		(0.155)		(0.207)		(0.206)
Skewness			-0.004***	-0.003**			-0.006***	-0.006***
			(0.001)	(0.001)			(0.002)	(0.002)
Flow	0.374***	0.370***	0.381***	0.379***	0.281***	0.280***	0.279***	0.278***
	(0.018)	(0.018)	(0.017)	(0.017)	(0.009)	(0.009)	(0.009)	(0.009)
Log(Size)	-0.144***	-0.161***	-0.117***	-0.121***	-0.060**	-0.060**	-0.059**	-0.060**
	(0.032)	(0.041)	(0.017)	(0.017)	(0.027)	(0.027)	(0.028)	(0.028)
Log(Age)	-0.523***	-0.501***	-0.534***	-0.531***	-0.701***	-0.705***	-0.712***	-0.716***
	(0.033)	(0.042)	(0.030)	(0.030)	(0.041)	(0.041)	(0.042)	(0.042)
Log(Fsize)	0.036**	0.045**	0.023**	0.027***	0.114***	0.118***	0.117***	0.121***
	(0.017)	(0.020)	(0.009)	(0.009)	(0.012)	(0.012)	(0.012)	(0.012)
BigFam	0.002***	0.002***	0.002***	0.002***	0.002**	0.001*	0.002*	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Star Fund	0.010***	0.010***	0.011***	0.011***	0.009***	0.009***	0.009***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Non12b-1	-0.438***	-0.431***	-0.438***	-0.426***	-0.308***	-0.294***	-0.310***	-0.295***
	(0.052)	(0.056)	(0.049)	(0.048)	(0.064)	(0.063)	(0.066)	(0.066)
12b-1	-0.635***	-0.659***	-0.605***	-0.607***	0.079	0.068	0.075	0.065
	(0.098)	(0.105)	(0.095)	(0.098)	(0.202)	(0.202)	(0.207)	(0.206)
Rear Load	-0.098***	-0.091**	-0.124***	-0.125***	-0.104***	-0.104***	-0.108***	-0.108***
	(0.031)	(0.037)	(0.014)	(0.014)	(0.036)	(0.036)	(0.036)	(0.037)
Front Load	-0.018	-0.014	-0.005	-0.005	-0.082***	-0.079**	-0.085***	-0.082***
	(0.013)	(0.010)	(0.005)	(0.005)	(0.031)	(0.031)	(0.031)	(0.031)
Turnover	-0.116	-0.106	-0.026	-0.017	-0.070***	-0.061**	-0.081***	-0.070***
	(0.089)	(0.088)	(0.020)	(0.019)	(0.025)	(0.024)	(0.025)	(0.024)
Perf	0.172***	0.193***	0.179***	0.203***	0.132***	0.141***	0.135***	0.143***
	(0.018)	(0.018)	(0.016)	(0.015)	(0.021)	(0.020)	(0.021)	(0.020)
(Perf)2	1.778***	1.926***	1.332***	1.447***	1.130***	1.174***	1.114***	1.198***
	(0.473)	(0.535)	(0.299)	(0.323)	(0.409)	(0.422)	(0.421)	(0.434)
N	435077	435074	434921	434918	191546	191546	183787	183787
R-sq	0.24	0.25	0.25	0.25	0.14	0.14	0.14	0.14

positive relation between MAX and future flows can be explained by a preference for lottery-like payoffs that retail investors are more likely to exhibit.

While there is a significant difference in the coefficients of MAX between retail and non-retail funds, the estimates of other variables are similar in both groups. Most importantly, the coefficient of style-adjusted returns in retail funds and non-retail funds are similar and the flow-performance relation is convex in both groups in all specifications. This finding suggests that the dynamics behind the flow-MAX relation cannot be attributed to the previously documented flow-performance relation. Age, size, and expenses negatively affect flows (Sirri and Tufano (1998), Chevalier and Ellison (1997)). Interestingly, retail investors seem to be less responsive to turnover of the fund as suggested by the insignificant coefficient, while there is a significant negative effect of turnover on future flow in non-retail funds.

Since retail funds are mainly bigger in size compared to non-retail funds, the finding that MAX-flow relation is more pronounced among retail funds might be an artifact of size effect.[36] In order to examine this possibility, I divide the sample into groups based on median size (i.e. above-median/below-median), and run separate regressions within these groups. All regressions include a dummy variable (retail) that is equal to 1 if the fund is a retail fund, zero otherwise, and an interaction term of MAX with this dummy. Other control variables are same as in Table 1.8. If retail fund effect is not subsumed by fund size, then the coefficient of the dummy variable should be positive. For brevity, I only report the coefficients of MAX, interaction term (MAX*retail),

[36] In my sample, retail funds have an average TNA of $569 million while the average TNA of non-retail funds is $270 million. The median TNA of both groups is around $52 million, suggesting that there is more skewness in retail fund size. It might be the case that there is a significant MAX-flow relation in relatively big non-retail funds, or no MAX-flow relation in small retail funds.

volatility, and skewness in Table 1.9. The interaction terms from the regressions in the above-median group are all highly significant, while the coefficients of MAX are not.[37] Moreover, I also find similar results in the below-median group. Hence, the results suggest that even after controlling for size, MAX-flow relation is confined to retail funds and do not exist in non-retail funds.

Interpreting the Results

My findings suggest that investors prefer mutual funds that exhibit extreme positive returns. If MAX is a predictor of risk-adjusted future performance, then selecting based on MAX may be simply an easy way for investors to select well-managed funds. But even if MAX-based fund investments underperform, their selection could be rational. For example, if lottery-like features enter investors' utility functions, then choosing such funds may be optimal if MAX predicts future MAX. In other words, if lottery-like features are persistent and valuable to investors, then MAX-based flows can be rational.

Preferences for extreme returns have been analyzed before. Using the cumulative prospect theory, Barberis and Huang (2008) show that investors, through a weighting function, overweight the probabilities in the extremes. For these investors, an asset that provide extreme payoff is desirable, even though the probability of the payoff is fairly small. Therefore, investors are willing to pay high prices for this asset and accept a lower return on average. In addition, Brunnermeier, Gollier, and Parker (2007) develop a model in which investors endogenously increase their beliefs about the probabilities of high payoff states in order to maximize their felicity, defined as the discounted present value of expected flow utilities. The authors show that this distortion in beliefs leads to a strong

[37] There is only one case where the coefficient of MAX is itself significant. However, this result is not robust to alternative specifications.

Table 1.9. Flow-MAX Relation in Retail/Non-Retail Funds after Controlling for Size

This table reports the parameter estimates from Fama and MacBeth (1973) regressions of monthly observations from 1998/10 to 2010/12 within different size groups. The dependent variable is the flow ratio. Retail is a dummy variable that takes the value of 1 if the fund is a retail fund, zero otherwise. Panel A presents the results for above median size group, whereas in Panel B the results for below median size group are presented. Other control variables (omitted) are same as in Table 1.8. Standard errors (in parenthesis) are corrected for heteroskedasticity and autocorrelation (up to 6 lags) using Newey and West (1987).

	Above Median							
	Piecewise Linear Regression				Quadratic Regression			
MAX	0.025	0.123*	0.028	0.127**	0.009	0.102	0.010	0.105*
	(0.052)	(0.064)	(0.053)	(0.063)	(0.054)	(0.064)	(0.055)	(0.063)
MAX*Retail	0.215***	0.160***	0.215***	0.158***	0.183***	0.125**	0.185***	0.125**
	(0.053)	(0.053)	(0.055)	(0.055)	(0.050)	(0.050)	(0.053)	(0.053)
Volatility		-0.491***		-0.504***		-0.500***		-0.514***
		(0.142)		(0.130)		(0.145)		(0.134)
Skewness			-0.005***	-0.004***			-0.005***	-0.004***
			(0.001)	(0.001)			(0.001)	(0.001)
N	323545	323545	319389	319389	323545	323545	319389	319389
R-sq	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31

	Below Median							
	Piecewise Linear Regression				Quadratic Regression			
MAX	-0.043	0.047	-0.031	0.055	-0.045	0.039	-0.037	0.045
	(0.061)	(0.067)	(0.060)	(0.065)	(0.061)	(0.067)	(0.060)	(0.066)
MAX*Retail	0.173***	0.159**	0.188***	0.176***	0.135**	0.124**	0.150**	0.140**
	(0.066)	(0.065)	(0.063)	(0.062)	(0.064)	(0.062)	(0.062)	(0.061)
Volatility		-0.526***		-0.548***		-0.608***		-0.640***
		(0.192)		(0.187)		(0.192)		(0.186)
Skewness			-0.004***	-0.003**			-0.004***	-0.003**
			(0.001)	(0.001)			(0.001)	(0.001)
N	303078	303075	299319	299319	303078	303075	299319	299319
R-sq	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18

preference toward assets with lottery-like features and a relatively lower level of average returns from investments.[38] My results are consistent with the predictions of both models. Given the preference of investors for lottery-like features in asset returns, high MAX

[38] In Barberis and Huang (2008), the overweighting of tails does not represent a bias in beliefs. It is exogenously determined by the weighting function and captures the preference for lottery-like return distribution. In Brunnermeier, Gollier and Parker (2007), investors choose to bias upward their beliefs about the likelihood of positive extreme states since they benefit from this optimism by increasing their utility. That is, biases in beliefs are determined endogenously by the economic environment.

funds may be attractive to investors if they anticipate similar extreme payoffs from these funds in the future. That is, once MAX is realized for a fund, investors may perceive that future extreme payoffs are more likely and their asset allocation decisions are affected by the *expectation* of this upside potential given their preferences.[39] This expectation is rational if MAX exhibits persistence. Hence, if MAX indeed exhibits persistence and investors derive utility from having a preference to extreme payoffs as in Brunnermeier, Gollier, and Parker (2007), then high MAX funds can maximize their utility. Therefore, understanding whether MAX predicts future returns or the likelihood of future extreme returns or both is important to understand the implications of my findings. In this section, I investigate both issues by analyzing the persistence and the future performance of high MAX funds.

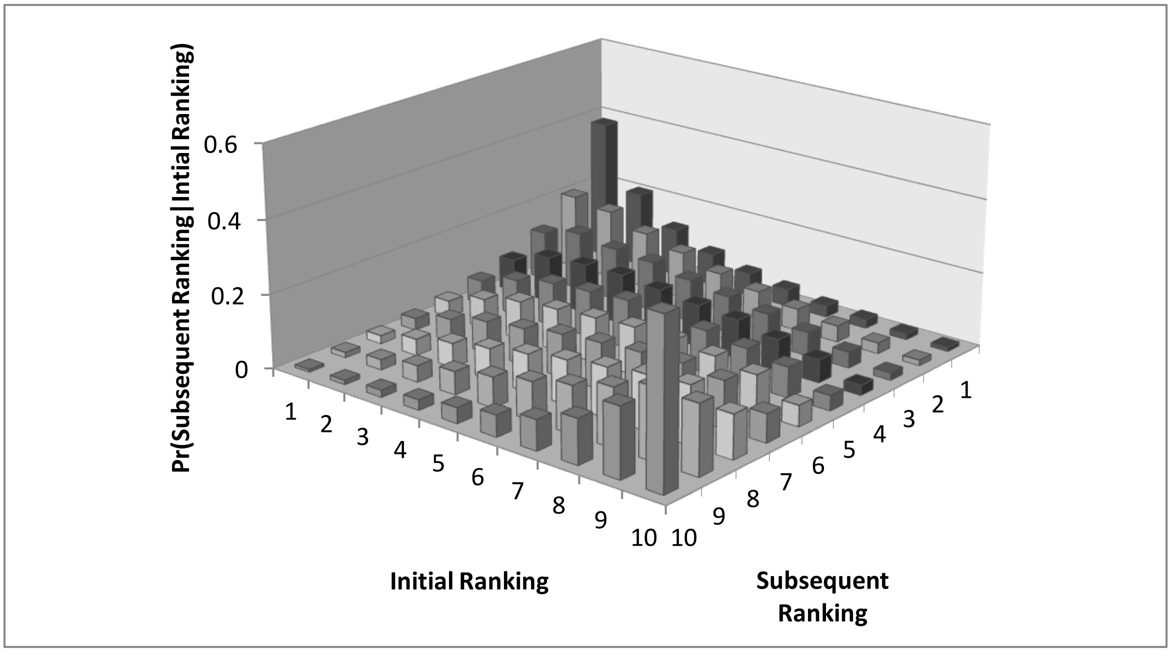
Persistence of MAX

To examine the persistence of MAX, I first form a transition matrix which gives the probability of achieving a ranking of decile j given an initial ranking of decile i over the previous month. The results are presented in Figure 1.1. If there is no persistence, then all probabilities should be approximately 10%. Instead, MAX exhibits clearly persistence especially in the top and bottom deciles. Funds in the top decile (in one of the top three deciles) with respect to MAX are in the top decile (in one of the three deciles) in the subsequent month with a probability of 49% (79%). Moreover, in untabulated results, I find that funds in the top decile with respect to MAX are in the top decile in month t+3, t+6, and t+12 with probabilities of 45%, 43%, and 39%. In an alternative test, for each month t+1, I regress MAX_{t+1} within that month onto MAX_t from the previous month and

[39] In both Barberis and Huang (2008) and Brunnermeier, Gollier and Parker (2007), investors overestimate the probabilities of *future* extreme payoff states.

Figure 1.1. Transition Matrix

Each month, funds are sorted into deciles according to MAX in that month. Independently, funds are sorted into deciles according to MAX in the subsequent month. The figure shows the average probability of a fund in decile i according to MAX in one month to be in decile j according to MAX in the subsequent month.



eight lagged (i.e. month t) control variables including the logarithm of size, the logarithm of age, Non12b-1 fees,12b-1 fees, front load, rear load, turnover, style-adjusted returns, volatility, and skewness. In addition, I conduct the same regressions using $MAX_{t+3,}$ MAX_{t+6}, and MAX_{t+12} as dependent variables. Table 1.10 reports parameter estimates. In the univariate regression of MAX_{t+1} on lagged MAX_t, the average cross-sectional coefficient is 0.629 with a Newey and West (1987) adjusted t-statistics of 38.08. The R^2 of the regression is 38%, which indicates a considerable explanatory power. When the eight control variables are added to the regression, the coefficient on MAX_t remains highly significant with a value of 0.612. When $MAX_{t+3,}$ MAX_{t+6}, and MAX_{t+12} are used as dependent variables, the coefficient of MAX_t is positive, large, and extremely statistically significant, suggesting that the persistence of MAX is also strong in the long run. Overall, both tests suggest that funds with high MAX in one month are more likely to show the same features in the future. If investors indeed derive utility from having a preference to extreme payoffs with small probabilities, then it is more likely to gain extreme payoffs if they follow high MAX funds.

Performance-MAX Relation

To maximize gains from an investment, investors should direct flows into funds with better future performance. The previous section shows that funds with high MAX in month t enjoy higher flows in month t+1 relative to other funds that have similar characteristics. One natural question to ask is whether the funds with high MAX also perform better in the future (i.e. in month t+2) compared to other funds that have similar characteristics. In other words, I investigate whether investors, who follow a MAX-based signal to select their fund investments in month t+1, experience better performance. To address this question, I use the following Fama-MacBeth's (1973) regression model:

$$\text{Performance}_{i,t+2} = \alpha + \beta_{i,t} \times \text{MAX}_{i,t} + \delta_{i,t} \times \text{X}_{i,t} + \varepsilon_{i,t} \qquad (1.4)$$

where $\text{Performance}_{i,t+2}$ refers to the performance of fund i from month t+2. I use five different performance measures: raw return, style-adjusted returns, risk-adjusted returns according to one factor (CAPM), Fama and French's (1993) three factors and Carhart's (1997) four factors model. To calculate the risk-adjusted returns, I proceed as follows: For every month t in my sample, I regress funds' excess returns on the risk factors over the previous 60 months (i.e. from month t-60 to month t-1) and save the factor loadings. I

Table 1.10. Cross sectional Predictability of MAX

This table reports the parameter estimates from Fama and MacBeth (1973) regressions of monthly observations from 1998/10 to 2010/12 for retail and non-retail funds. The dependent variable is MAX in month t+1, t+3, t+6 or t+12. The independent variables are MAX in month t and a set of control variables including style-adjusted returns, flows, the logarithm of size, the logarithm of age, 12b-1 fees, Non12b-1 fees (as measured by the difference between expense ratios and 12b-1 fees), rear loads, front load, turnover, volatility and skewness. All control variables are lagged at least one month. Standard errors (in parenthesis) are corrected for heteroskedasticity and autocorrelation (up to 6 lags) using Newey and West (1987).

Dep. Variable	MAX_{t+1}		MAX_{t+3}		MAX_{t+6}		MAX_{t+12}	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MAX	0.629***	0.612***	0.630***	0.579***	0.604***	0.550***	0.553***	0.521***
	(0.017)	(0.018)	(0.022)	(0.019)	(0.028)	(0.029)	(0.034)	(0.040)
Return		-0.042***		-0.040***		-0.034***		-0.035***
		(0.003)		(0.004)		(0.005)		(0.004)
Flow		0.001***		0.001***		0.001***		0.001***
		(0.000)		(0.000)		(0.000)		(0.000)
Log(Size)		0.012		0.024**		0.027**		0.031*
		(0.012)		(0.011)		(0.012)		(0.016)
Log(Age)		0.019		0.016		0.030		0.035
		(0.020)		(0.024)		(0.029)		(0.034)
Non12b-1		0.124***		0.142***		0.156***		0.175***
		(0.013)		(0.015)		(0.016)		(0.019)
12b-1		-0.020***		-0.020***		-0.017***		-0.019***
		(0.002)		(0.002)		(0.002)		(0.003)
Rear Load		-0.002***		-0.003***		-0.003***		-0.004***
		(0.001)		(0.001)		(0.001)		(0.001)
Front Load		-0.004***		-0.004***		-0.004***		-0.005***
		(0.000)		(0.001)		(0.001)		(0.001)
Volatility		0.144***		0.116***		0.084**		0.022
		(0.028)		(0.030)		(0.040)		(0.040)
Skewness		0.001**		0.001*		0.000		-0.001*
		(0.000)		(0.000)		(0.000)		(0.000)
N	719883	652864	782175	634782	742055	608183	665931	556237
R-sq	0.38	0.46	0.38	0.42	0.36	0.38	0.31	0.33

require at least 48 months in the preceding 60 months to run the regression.[40] Then, I estimate the risk-adjusted return in month t as the difference between the fund's realized

[40] Although a longer estimation period excludes a greater fraction of funds from the sample, this is not a major problem for my sample since I utilize monthly return data before 1998. If a fund has monthly returns data before October, 1998- the first month my sample starts- then, I use all available data in order to calculate risk-adjusted returns. This allows me to keep as many data points as possible for each fund after

excess return and the required return, defined as the factor loadings times the corresponding factor realizations in month t. The control variables include the logarithm of fund size, the logarithm of fund age, Non12b-1, 12b-1, turnover, monthly flow, front load, rear load, and performance (Chen et. al. (2004)). All control variables are lagged at least one month.

Results are in Table 1.11. The coefficients of MAX are negative in all specifications. An increase of 1% daily return in MAX decreases raw returns by 10 bps and style-adjusted returns by 9 bps in month t+2. For risk-adjusted returns, the effect of MAX varies between 7.7 bps and 13.5 bps. There is some evidence of short-term performance persistence as revealed by the positive coefficient of the lagged performance. That is, even though both high MAX and better performance lead to higher flows in the future, the effect of MAX on future performance is different than that of prior performance.[41]

I also use 3-month (from t+2 to t+4), 6-month (from t+2 to t+7), and 12-month (from t+2 to t+13) returns as dependent variables in order to analyze the long-term relation between MAX and performance. The regressions use the same control variables. In Table 1.12, I only report the coefficients for MAX and omit the coefficients for other control variables from the table. All coefficient estimates are negative and in all but one specification they are statistically significant at conventional levels. The coefficients vary

---

October, 1998. However, the results are qualitatively similar if I do not use data before 1998, and hence start the sample by 2003. Moreover, using a longer estimation period also reduces the sampling error in betas and mitigates the effect of incubation bias that affects mostly the subset of young funds in the CRSP database. Nevertheless, I also estimate the betas using previous 36 months and the results are qualitatively similar.

[41] In further tests, I control for volatility and skewness. Moreover, I also run a regression with flows andreturns in month t+1 as control variables. Results are similar and available upon request.

Table 1.11. Flow-Performance Relation

This table reports the parameter estimates from Fama and MacBeth (1973) regressions of monthly observations for funds from 1998/10 to 2010/12. The dependent variable is the fund performance in month t+2. The independent variables are MAX and a set of control variables including flows, the logarithm of size, the logarithm of age, 12b-1 fees, Non12b-1 fees (as measured by the difference between expense ratios and 12b-1 fees), rear loads, front load, turnover and fund performance. All independent variables are lagged at least two months and all specifications include style-dummies. I use five different performance measures: raw return, style-adjusted return, and risk-adjusted returns according to one factor (CAPM), Fama and French's (1993) three factors and Carhart's (1997) four-factor models. Standard errors (in parenthesis) are corrected for heteroskedasticity and autocorrelation (up to 6 lags) using Newey and West (1987).

	Raw Return	Style-adjusted Return	Capm-adjusted Return	FF-adjusted Return	Carhart-adjusted Return
MAX	-0.105***	-0.091*	-0.135**	-0.107**	-0.077**
	(0.038)	(0.054)	(0.052)	(0.043)	(0.037)
Flow	0.003*	0.003*	0.003*	0.004**	0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Log(Size)	-0.007	-0.012	-0.002	0.05	0.042
	(0.044)	(0.051)	(0.045)	(0.056)	(0.064)
Log(Age)	-0.045	-0.03	0.042	-0.073	-0.056
	(0.097)	(0.112)	(0.103)	(0.116)	(0.110)
Non12b-1	-0.048**	-0.040*	-0.072**	-0.078**	-0.067*
	(0.021)	(0.024)	(0.030)	(0.031)	(0.035)
12b-1	-0.095***	-0.099***	-0.090***	-0.077***	-0.079***
	(0.015)	(0.017)	(0.019)	(0.015)	(0.014)
Rear Load	-0.001	-0.002	-0.001	-0.003	-0.003
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Front Load	-0.004**	-0.004**	-0.003	-0.004**	-0.003*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Turnover	-0.032	0.110	0.063	0.101	-0.166
	(0.360)	(0.407)	(0.377)	(0.322)	(0.300)
Perf	0.040*	0.042*	0.043*	0.032	0.026
	(0.023)	(0.025)	(0.022)	(0.019)	(0.017)
N	639545	639545	461539	461539	461539
R-sq	0.42	0.12	0.40	0.22	0.20

between 23 bps and 35 bps for 3-month returns, 36 bps and 56 bps for 6-month returns, and 44 bps and 87 bps for 12-month returns. To sum up, these results show that there is a negative and significant relationship between future fund performance and MAX, which suggests that directing incremental flows into funds with high MAX is costly for fund investors in terms of expected returns. This finding is consistent with Frazzini and

Lamont (2008), who argue that retail investors direct their money to funds with overpriced stocks, and with Bali, Cakici, and Whitelaw (2011) that show that stocks with extreme positive returns are overpriced. On the other hand, Edelen (1999) states that fund managers engage in costly liquidity based trading in order to provide liquidity services to their investors. Edelen (1999) estimates that around 28% of total trading activity can be characterized as liquidity motivated and is associated with an estimated 1.5% decline in abnormal returns. Alexander, Cici, and Gibson (2007) show that liquidity trades occur mostly during fund inflows while fund outflow cause valuation motivated trades. Hence if MAX leads to more inflows then it may also cause more liquidity based trading and the subsequent underperformance.

Table 1.12. Flow-Performance Relation over Longer Horizons

This table reports the parameter estimates of MAX from Fama and MacBeth (1973) regressions of monthly observations for funds from 1998/10 to 2009/12. The dependent variables are 3-month (from t+2 to t+4), 6-month (from t+2 to t+7), and 12-month fund performance. The independent variables are MAX and a set of control variables including flows, the logarithm of size, the logarithm of age, 12b-1 fees, Non12b-1 fees (as measured by the difference between expense ratios and 12b-1 fees), rear loads, front load, turnover and fund performance. All independent variables are lagged at least two months and all specifications include style-dummies. I use five different performance measures: raw return, style-adjusted return, and risk-adjusted returns according to one factor (CAPM), Fama and French's (1993) three factors and Carhart's (1997) four-factor models. Standard errors (in parenthesis) are corrected for heteroskedasticity and autocorrelation (up to 6 lags) using Newey and West (1987).

	Raw Return	Style-adjusted Return	Capm-adjusted Return	FF-adjusted Return	Carhart-adjusted Return
3-month	-0.248**	-0.233*	-0.353***	-0.297***	-0.238**
	(0.104)	(0.135)	(0.125)	(0.099)	(0.093)
6-month	-0.385*	-0.358*	-0.559**	-0.491***	-0.354**
	(0.179)	(0.206)	(0.217)	(0.163)	(0.160)
12-month	-0.519**	-0.443	-0.870***	-0.738***	-0.486**
	(0.327)	(0.336)	(0.365)	(0.284)	(0.222)

A Closer Look at Idiosyncratic Volatility, Skewness and MIN

In previous tests, I use total volatility and skewness as control variables. Since MAX is almost exclusively idiosyncratic in nature, there may be still a concern whether MAX is a good proxy for idiosyncratic volatility or skewness rather than total volatility or skewness.[42] Moreover, similar to the attraction to high MAX funds, investors may dislike funds with minimum extreme returns (MIN).[43] Therefore, to assess the robustness of my results, I analyze the link, if any, between idiosyncratic volatility, idiosyncratic skewness and MIN.

Idiosyncratic volatility (Ivolatility) is calculated by the standard deviation of the residuals from a single factor in each month t:

$$\mathrm{R}_{i,d} - \mathrm{r}_{f,d} = \alpha + \beta_i \times (\mathrm{R}_{m,d} - \mathrm{r}_{f,d}) + \varepsilon_{i,d} \tag{1.5}$$

where $\mathrm{R}_{i,d}$, $\mathrm{R}_{m,d}$, $\varepsilon_{i,d}$ are the fund return, market returns and the idiosyncratic return on day d.[44] I follow Harvey and Siddique (2000) in order to decompose skewness into idiosyncratic and systematic components by estimating the following regression for each fund over the previous 12 months:

$$\mathrm{R}_{i,d} - \mathrm{r}_{f,d} = \alpha + \beta_i \times (\mathrm{R}_{m,d} - \mathrm{r}_{f,d}) + \theta_i \times (\mathrm{R}_{m,d} - \mathrm{r}_{f,d})^2 + \varepsilon_{i,d} \tag{1.6}$$

The idiosyncratic skewness (Iskewness) is defined as the skewness of daily residuals $\varepsilon_{i,d}$. Lastly, analogous to MAX, MIN is defined as the style-adjusted minimum daily returns

[42] Mitton and Vorkink (2007) develop a model in which investors have heterogeneous preference for skewness. The model demonstrates that, given the lack of diversification in investor holdings, investors overvalue the stocks with higher idiosyncratic skewness and earn lower expected returns from these stocks.

[43] The theoretical work of Barberis and Huang (2008) predicts that investors overweight the probabilities of large losses, and thus the effect of minimum daily returns (MIN) should be opposite of MAX. However, in the model of Brunnermeir, Gollier, and Parker (2007), investors underweight rather than overweight minimum returns, and hence the pricing implication of minimum returns is different than that of Barberis and Huang (2008).

[44] The results hold if I use the four factor model instead.

over a month. Not surprisingly, MAX is highly correlated with Ivolatility and MIN. The average cross-sectional correlation between MAX and Ivolatility and MAX 70%, while the correlation between MAX and MIN is -66%. On the other hand, the correlation between MAX and Iskewness is very low (0.6%).

I first construct double sorted portfolios. That is, I sort all funds into three groups-Bottom, Mid, Top-based upon a 20:60:20 split. Then within each group, I further sort the funds into deciles based on MAX within the same month. This creates 30 portfolios which are held for a month after portfolio formation. Table 1.13 reports average flows for 30 portfolios in the subsequent month and the flow difference between high MAX (decile 10) and low MAX (decile 1) portfolios. The results suggest that the effect of MAX is robust to controlling for Ivolatility, Iskewness and MIN. The flow difference between High and Low MAX portfolios is greater for funds with higher Ivolatility, lower Iskewness, and lower MIN. Next, I conduct Fama and MacBeth's (1973) regressions using the whole sample as well as within the retail and non-retail funds. Table 1.14 reports the time series average of cross-sectional regressions. In the regressions with the whole sample, the coefficients of MAX range between 0.251 and 0.094, and are highly statistically significant. Idiosyncratic volatility and idiosyncratic skewnesss affect fund flows negatively, echoing the regression estimates of total volatility and skewness. The effect of MIN is also negative, albeit weaker than that of MAX, suggesting that the response of investors to extreme returns is asymmetric. It might be due to the fact that mostly investors that already purchased a fund would react to MIN, while MAX may attract both the existing and potential investors in the market. When I divide the whole

Table 1.13. Average Flow of Funds Sorted on MAX after Controlling for Idiosyncratic Volatility, Idiosyncratic Skewness, and MIN

Each month t from 1998/9 to 2010/12, funds are ranked in ascending order according to idiosyncratic volatility, idiosyncratic skewness or MIN in month t and are allocated into three groups- Bottom, Mid, Top-based upon a 20:60:20 split. Bottom (Top) refers to the group of funds that are in the bottom (top) 20% of all funds ranked according to the fund characteristic. Mid includes funds that are in the 2nd-4th quintiles according to the fund characteristics. Then within each group, I further sort the funds into deciles based on MAX within the same month. This creates 30 portfolios which are held for a month after portfolio formation. This table reports equally weighted monthly portfolio flows (for each decile in month t+1. Decile 1(10) is the group of funds with the lowest (highest) MAX. Flow is the growth rate of assets under management due to new investments. The last row (HM-LM) presents the difference of flow ratios between the High MAX (decile 10) and Low MAX (decile 1) groups. Newey and West (1987) adjusted t-statistics (up to 6 lags) are reported in parenthesis. All values except the t-statistics are given in percentage terms.

	Iskewness			Ivolatility			Min		
	Bottom	Mid	Top	Bottom	Mid	Top	Bottom	Mid	Top
Low MAX	0.49	0.35	0.12	0.38	0.33	0.29	0.22	0.26	0.41
2	0.73	0.46	0.08	0.30	0.43	0.39	0.36	0.35	0.36
3	0.77	0.40	0.23	0.26	0.42	0.56	0.44	0.43	0.36
4	0.81	0.50	0.19	0.30	0.53	0.62	0.47	0.51	0.47
5	0.83	0.58	0.18	0.45	0.60	0.67	0.59	0.53	0.62
6	0.98	0.58	0.26	0.42	0.59	0.81	0.63	0.58	0.43
7	0.97	0.56	0.41	0.41	0.63	0.85	0.80	0.61	0.52
8	1.12	0.54	0.31	0.51	0.62	1.03	0.95	0.64	0.62
9	1.17	0.68	0.56	0.61	0.71	1.00	0.99	0.76	0.79
High MAX	1.44	1.03	0.83	0.89	0.93	1.38	1.49	1.01	0.94
HM-LM	0.95	0.68	0.71	0.50	0.61	1.10	1.27	0.75	0.53
	(4.23)	***(4.48)***	***(3.38)***	***(4.41)***	***(4.85)***	***(5.71)***	***(6.17)***	***(4.39)***	***(3.61)***

the sample into retail and non-retail funds, I find that the coefficient of MAX is significant only in retail funds. On the other hand, the relation between fund flows and other fund characteristics, such as performance, holds for both groups. Overall, these additional tests suggest that the impact of MAX is robust to using idiosyncratic volatility, skewness, and MIN.

Some Robustness Tests

In my previous analysis, I control for volatility and skewness to show that the effect of MAX is not subsumed by these variables. I measure volatility using daily

Table 1.14. Flow-MAX Relation with Ivolatility, Iskewness, and MIN

This table reports the parameter estimates from Fama and MacBeth (1973) regressions of monthly observations from 1998/10 to 2010/12 for all funds as well as retail and non-retail funds. The dependent variable is the flow ratio. The independent variables are MAX over the previous month and a set of control variables including flows, the logarithm of size, the logarithm of age, 12b-1 fees, Non12b-1 fees (as measured by the difference between expense ratios and 12b-1 fees), rear loads, front load, turnover and fund performance. All control variables are lagged at least one month. All specifications include style-dummies. Standard errors (in parenthesis) are corrected for heteroskedasticity and autocorrelation (up to 6 lags) using Newey and West (1987).

	All Funds			Retail Funds			Non-Retail Funds		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
MAX	0.251***	0.139***	0.094**	0.201**	0.158***	0.160**	0.172*	0.078	0.098
	(0.050)	(0.027)	(0.037)	(0.077)	(0.030)	(0.065)	(0.090)	(0.053)	(0.066)
Ivolatility	-0.468***			-0.127			-0.405**		
	(0.161)			(0.317)			(0.193)		
Iskewness		0.000			0.000			0.001	
		0.000			0.000			0.000	
MIN			-0.070*			-0.049			0.047
			(0.041)			(0.058)			(0.067)
Flow	0.346***	0.347***	0.346***	0.372***	0.382***	0.373***	0.281***	0.280***	0.281***
	(0.014)	(0.014)	(0.014)	(0.018)	(0.017)	(0.018)	(0.009)	(0.009)	(0.009)
Log(Size)	-0.110***	-0.109***	-0.110***	-0.156***	-0.118***	-0.157***	-0.061**	-0.058**	-0.061**
	(0.019)	(0.019)	(0.019)	(0.039)	(0.017)	(0.039)	(0.027)	(0.028)	(0.027)
Log(Fsize)	0.042***	0.044***	0.044***	0.037**	0.026***	0.042**	0.114***	0.118***	0.115***
	(0.007)	(0.008)	(0.008)	(0.016)	(0.009)	(0.018)	(0.012)	(0.012)	(0.012)
Big Family	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002*	0.002*	0.002*
	0.000	0.000	0.000	0.000	0.000	0.000	(0.001)	(0.001)	(0.001)
Star Fund	0.011***	0.011***	0.011***	0.011***	0.011***	0.011***	0.009***	0.009***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Log(Age)	-0.580***	-0.586***	-0.581***	-0.502***	-0.535***	-0.494***	-0.701***	-0.710***	-0.703***
	(0.029)	(0.030)	(0.029)	(0.043)	(0.030)	(0.050)	(0.041)	(0.042)	(0.041)
Non12b-1	-0.391***	-0.406***	-0.404***	-0.470***	-0.436***	-0.428***	-0.299***	-0.312***	-0.306***
	(0.045)	(0.046)	(0.046)	(0.065)	(0.049)	(0.057)	(0.065)	(0.065)	(0.065)

Table 1.14 (Continued)

	All Funds			Retail Funds			Non-Retail Funds		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
12b-1	-0.707***	-0.736***	-0.710***	-0.596***	-0.610***	-0.630***	0.092	0.097	0.083
	(0.104)	(0.101)	(0.103)	(0.095)	(0.095)	(0.097)	(0.203)	(0.209)	(0.203)
Rear Load	-0.136***	-0.140***	-0.135***	-0.095***	-0.124***	-0.095***	-0.102***	-0.106***	-0.101***
	(0.016)	(0.015)	(0.016)	(0.033)	(0.014)	(0.034)	(0.036)	(0.037)	(0.036)
Front Load	-0.023***	-0.025***	-0.023***	-0.017	-0.006	-0.014	-0.082***	-0.087***	-0.081***
	(0.006)	(0.005)	(0.006)	(0.012)	(0.005)	(0.010)	(0.031)	(0.032)	(0.031)
Turnover	-0.060**	-0.057**	-0.065**	-0.089	-0.037	-0.145	-0.089***	-0.084***	-0.091***
	(0.026)	(0.027)	(0.027)	(0.054)	(0.028)	(0.094)	(0.027)	(0.027)	(0.027)
Low	0.011***	0.012***	0.014***	0.015***	0.013***	0.011**	0.005	0.005	0.006*
	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)	(0.005)	(0.003)	(0.003)	(0.003)
Mid	0.004***	0.005***	0.005***	0.004***	0.005***	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
High	0.033***	0.033***	0.033***	0.038***	0.034***	0.039***	0.023***	0.023***	0.023***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.004)	(0.004)	(0.004)	(0.004)
N	626648	618736	626651	435097	434944	435100	191551	183792	191551
R-sq	0.21	0.21	0.21	0.24	0.25	0.24	0.14	0.14	0.14

returns over a month, while skewness is measured using daily returns over the previous 12 months. Even though these measurement periods are the mostly commonly used in the existing literature, in this section, I use alternative measurement periods to assess the robustness of my results. Specifically, I calculate volatility using daily returns over the previous 3,6, and 12 months, and skewness using daily returns over the previous 1, 3, and 6 months. Panel A of Table 1.15 reports the coefficients of MAX from piecewise linear regressions as in Table 1.6, when volatility and skewness with alternative measurement periods are used.[45] The effect of MAX in all regressions is highly statistically significant, and if anything, the effect of MAX increases as skewness measured over shorter periods is used.

In Table 1.4, I sort funds according to alternative definitions of MAX to see whether there is a flow spread among fund portfolios when alternative MAX measures are used in sorting. In this section, I test the flow-MAX relation in a multivariate setting by using these alternative definitions of MAX. Specifically, I use: (1) MAX measured as the average of highest 2, 3, 4, 5 style-adjusted daily returns over a month, (2) MAX measured as the single highest style-adjusted daily return over the previous 3, 6, and 12 months, (3) the average of single day MAX in a month over the previous 3, 6, and 12 months. For brevity, Panel B of Table 1.15 only reports the coefficients of MAX both from piecewise linear and quadratic regressions. Three main results emerge: (1) The effect of MAX increases as I average over more days, (2) the effect of MAX declines but remains significant as longer periods are used, (3) the effect of MAX is substantially

[45] The results from quadratic regressions are similar and omitted for brevity.

Table 1.15. Flow-MAX Relation with Alternative Specifications

This table reports the parameter estimates from Fama and MacBeth (1973) regressions of monthly observations for funds from 1998/10 to 2010/12. Panel A reports the coefficient estimates of MAX from piecewise linear regressions when skewness and volatility are measured over alternative periods. Panel B report the coefficient estimates of MAX from piecewise and quadratic regressions when alternative MAX definitions are used. MAX [N] is the average of N (N=2, 3, 4, 5) highest daily style-adjusted returns in a month. MAX $[1]_K$ is the average highest daily style-adjusted return over the previous K months where K=3,6,12 months. $AMAX_3$ to $AMAX_{12}$ are the average of single day MAX over the previous 3, 6, 12 months.

Panel A: Coefficients of MAX with alternative measures of skewness and volatility measurement periods

		Skewness			
		1-month	3-month	6-month	12-month
Volatility	1-month	0.247***	0.249***	0.247***	0.239***
		(0.041)	(0.041)	(0.039)	(0.039)
	3-month	0.241***	0.230***	0.230***	0.223***
		(0.042)	(0.042)	(0.039)	(0.039)
	6-month	0.237***	0.229***	0.227***	0.219***
		(0.041)	(0.041)	(0.039)	(0.039)
	12-month	0.235***	0.224***	0.224***	0.212***
		(0.041)	(0.041)	(0.039)	(0.038)

Panel B: Coefficients of MAX with alternative MAX definitions

Piecewise Linear Regressions									
MAX[2]	MAX[3]	MAX[4]	MAX[5]	$MAX[1]_3$	$MAX[1]_6$	$MAX[1]_{12}$	$AMAX_3$	$AMAX_6$	$AMAX_{12}$
0.210***	0.258***	0.307***	0.349***	0.167***	0.129***	0.080***	0.333***	0.394***	0.417***
(0.035)	(0.042)	(0.048)	(0.055)	(0.025)	(0.024)	(0.021)	(0.038)	(0.046)	(0.048)

Quadratic Regressions									
MAX[2]	MAX[3]	MAX[4]	MAX[5]	$MAX[1]_3$	$MAX[1]_6$	$MAX[1]_{12}$	$AMAX_3$	$AMAX_6$	$AMAX_{12}$
0.166***	0.206***	0.250***	0.286***	0.163***	0.126***	0.074***	0.332***	0.404***	0.429***
(0.038)	(0.045)	(0.051)	(0.058)	(0.027)	(0.026)	(0.022)	(0.040)	(0.051)	(0.055)

greater if I use the average single day MAX in a month over past periods. The findings suggest that investors are increasingly attracted to MAX funds if these funds produce high MAX on average.

In my second essay, using a conditional setting, I show that long-term performance affects the sensitivity of flows to recent performance. Specifically, good (bad) long-term performance increases (decreases) the sensitivity of flows to recent performance. Hence, it is difficult to properly interpret the relation between short-term performance and flows from a model that omits long-term performance. In order to test whether the effect of MAX is robust to the presence of long-term performance, I calculate the average monthly returns over the previous 12 months (from t-1 to t-12)[46]. Then I run Fama and Macbeth's (1973) regressions that includes 12 months average performance as well as the interaction terms of these variables with short-term performance- Low, Mid, High- which is measured in month t. The dependent variable is flows in month t+1 and all other control variables are same as in Table 1.6. Results are shown in Table 1.16. Even though long-term performance significantly affects future flows, the coefficient of MAX also remains significant in all specifications, and ranges between 0.118 and 0.276.

Consistent with the results in my second essay, long-term performance positively affects the future flows as well as the sensitivity of flows to recent month performance. My results show that the effect of MAX and long-term performance do not subsume each other, but coexist. This is consistent with the idea that investors optimize their decisions based on a multi-attribute model, in which investors attribute positive value to different

[46] Using 24 or 36 months does not change inferences. Moreover, using average of risk-adjusted returns over the previous 12 months (from t-1 to t-12) does not change the inferences.

Table 1.16. Flow-MAX Relation after Controlling for Long-term Performance

This table reports the parameter estimates from Fama and MacBeth (1973) regressions of monthly observations from 1998/10 to 2010/12 for all funds. The dependent variable is the flow ratio. LTPerf is the average month returns over the previous 12 month from t-1 to t-12. The independent variables are MAX over the previous month and a set of control variables (omitted for brevity) including flows, the logarithm of size, the logarithm of age, 12b-1 fees, Non12b-1 fees (as measured by the difference between expense ratios and 12b-1 fees), rear loads, front load, turnover and fund performance. All control variables are lagged at least one month. All specifications include style-dummies. Standard errors (in parenthesis) are corrected for heteroskedasticity and autocorrelation (up to 6 lags) using Newey and West (1987).

Dep. Variables	Flows (in month t+1)		
MAX	0.118***	0.276***	0.184***
	(0.037)	(0.039)	(0.040)
Volatility		-0.627***	
		(0.149)	
Skewness			-0.000
			(0.001)
LTPerf	0.705***	0.621***	0.611***
	(0.092)	(0.069)	(0.075)
Low	0.015**	0.005	0.003
	(0.006)	(0.007)	(0.007)
Mid	0.001	0.002**	0.002**
	(0.001)	(0.001)	(0.001)
High	0.020***	0.025***	0.023***
	(0.006)	(0.005)	(0.005)
Low*LTPerf	1.098***	2.218***	2.290***
	(0.269)	(0.496)	(0.483)
Mid*LTPerf	0.176***	0.172***	0.162***
	(0.058)	(0.059)	(0.060)
High*LTPerf	0.945**	0.715*	0.765**
	(0.438)	(0.381)	(0.373)

fund characteristics (Fishbein and Azjen (1975), Lanchaster (1966), and Capon, Fitzsimons, and Prince (1996)). It might be the case that some investors care more about the long-term performance while the others follow a short-term approach and consider MAX as an important factor in their decisions. Since my data is not at the individual level, my tests do not allow differentiating among investors who use long-term performance, MAX or both. Nevertheless, if a sizable portion of investors use long-term performance or MAX in their decisions, then fund flows will reflect the effect of both characteristics.

Even though I correct standard errors for autocorrelation and heteroskedasticity, there might be still concern that standard errors are underestimated. Therefore, as a last check, I test the flow-Max relation in Table 1.17 by using panel regressions, where standard errors are clustered by fund and month (Peterson (2009)). Table 1.17 shows the results. In all specifications, MAX load positively and significantly, and its effect increase as I include volatility and skewness as additional control variables. One difference is that skewness has a negative but insignificant effect on fund flows while Fama-Macbeth's (1973) regressions show a negative and significant coefficient.[47] Overall, the effect of MAX on future flows does not depend on the regression method used.

Chapter Conclusion

Existing theoretical models argue that investors' trading decisions may be influenced by a preference for lottery-like features (i.e. high payoffs with small

[47] I also consider fixed-effect regressions with robust standard errors. Fixed effect regressions focus on the within-fund variation and ignore the between-fund variation. The result is robust to this specification as the coefficient of MAX is significant with a t-value of 2.35, suggesting that there is some variation in MAX within each fund.

Table 1.17. Flow-MAX Relation (Panel Regressions)

This table reports the parameter estimates from Fama and MacBeth (1973) regressions for funds from 1998/10 to 2010/12. The dependent variable is the flow ratio. The independent variables are MAX over the previous month and a set of control variables including volatility and skewness of fund returns, flows, the logarithm of size, the logarithm of age, the logarithm of family size, 12b-1 fees, Non12b-1 fees (as measured by the difference between expense ratios and 12b-1 fees), rear loads, front load, turnover, star-fund dummy, big-family dummy, and fund performance. All control variables are lagged at least one month and all specifications include style-dummies. The specifications in column 1, 3, and 5 use the normalized performance (Low, Mid, High), while the specifications in column 2, 4, and 6 use the style-adjusted return and its quadratic term. Standard are clustered by month.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MAX	0.118**	0.204***	0.125**	0.212***	0.104**	0.190***	0.111**	0.198***
	(0.050)	(0.046)	(0.052)	(0.048)	(0.049)	(0.045)	(0.051)	(0.046)
Volatility		-0.122**		-0.123**		-0.122**		-0.123**
		(0.049)		(0.050)		(0.050)		(0.051)
Skewness			-0.001	-0.001			-0.001	-0.001
			(0.001)	(0.001)			(0.001)	(0.001)
Flow	0.362***	0.361***	0.361***	0.360***	0.361***	0.361***	0.361***	0.360***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Log(Size)	-0.046*	-0.050**	-0.045*	-0.049**	-0.047**	-0.051**	-0.045*	-0.050**
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Log(Age)	-0.735***	-0.727***	-0.742***	-0.732***	-0.732***	-0.724***	-0.739***	-0.730***
	(0.035)	(0.034)	(0.035)	(0.034)	(0.035)	(0.034)	(0.035)	(0.034)
Log(Fsize)	0.010	0.014	0.009	0.013	0.010	0.013	0.008	0.013
	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)
BigFam	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Star Fund	0.009***	0.009***	0.009***	0.009***	0.010***	0.010***	0.010***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Non12b-1	-0.372***	-0.391***	-0.367***	-0.388***	-0.360***	-0.382***	-0.355***	-0.378***
	(0.048)	(0.046)	(0.048)	(0.046)	(0.048)	(0.047)	(0.048)	(0.047)
12b-1	-0.667***	-0.675***	-0.711***	-0.719***	-0.677***	-0.685***	-0.720***	-0.729***
	(0.075)	(0.075)	(0.070)	(0.070)	(0.074)	(0.075)	(0.070)	(0.070)
Rear Load	-0.115***	-0.116***	-0.117***	-0.118***	-0.115***	-0.116***	-0.117***	-0.118***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Front Load	-0.010	-0.011	-0.014*	-0.015**	-0.010	-0.011	-0.014*	-0.015**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Turnover	-0.018	-0.016	-0.020	-0.017	-0.018	-0.016	-0.020	-0.017
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Low	0.012***	0.012***	0.012***	0.012***				
	(0.002)	(0.002)	(0.002)	(0.002)				
Mid	0.004***	0.004***	0.004***	0.004***				
	(0.000)	(0.000)	(0.000)	(0.000)				
High	0.031***	0.029***	0.031***	0.029***				
	(0.003)	(0.003)	(0.003)	(0.003)				
Perf					0.116***	0.110***	0.116***	0.111***
					(0.012)	(0.012)	(0.012)	(0.012)
(Perf)2					0.289***	0.269***	0.292***	0.271***
					(0.096)	(0.094)	(0.097)	(0.095)
N	626623	626620	618708	618705	626623	626620	618708	618705
R-sq	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18

probabilities) in asset returns. Consistent with these arguments, I document a positive relationship between maximum style-adjusted daily returns (MAX) and future fund flows. My results are robust to various control variables including fund performance, fund size, age, turnover, fund fees, (idiosyncratic) volatility, and (idiosyncratic) skewness of fund returns. The flow-MAX relationship exists only among funds that cater primarily to retail investors. The results suggest that mutual fund investors, in particular retail investors, have a preference for funds with extreme positive returns although the probability of occurrence of large payoffs might be small. Because MAX is persistent, my results suggest that investors are successful in identifying lottery-like payoffs. But funds with high MAX underperform otherwise similar funds. This implies that selecting lottery like payoffs is costly.

My results have implications for the literature on determinants of mutual fund flows. While the question whether mutual fund managers strategically choose to have extreme daily returns is beyond the scope of this chapter, my results suggest that they would benefit from MAX as their compensation is tied to total assets under management. It would be interesting to see if fund managers have incentives to compete on MAX-based strategies. For instance, since high MAX funds underperform in the future, there will be a loss of flows. However, the additional flows from lottery-like returns may be high enough to compensate this cost, which might incentivize managers to take skewed-bets in order to boost their flows.

CHAPTER III

FUND FLOWS AND PERFORMANCE: LONG-TERM VS. SHORT-TERM

Introduction

Investors tend to buy funds with recent superior performance (Gruber (1996), Sirri and Tufano (1998)). Skilled managers are more likely to deliver higher performance and hence investors rationally flock to funds with recent superior performance assuming that it will persist into the future. These findings are consistent with theoretical models in the mutual fund literature, which generally posits that investors can infer the managerial ability from past returns (see Ippolito (1992), and Berk and Green (2004)) and direct capital to good performers. Investors can trade on recent signals particularly if there is imperfect information about the managers' ability (Scharfstein and Stein (1990)) or if managers' abilities and level of effort change in a dynamic environment (Ippolito (1992)). Accordingly, Sirri and Tufano (1998), Chevalier and Ellison (1997) show that prior year returns affect flows in the subsequent year. Cashman et. al. (2006) show a similar relation between monthly flows and prior month returns.

In this essay, I investigate the effect of long-term performance and short-term performance on future fund flows. Presumably there is less noisy and more information about manager skills in returns measured over a longer period compared to returns measured over a shorter period. For instance, a favorable quarter after a good two-year performance might be a stronger signal of managerial ability then a favorable quarter preceded by a poor two-year performance. More generally, the longer-term performance history may enhance the confidence with which investors assess the manager's ability. If so, this would affect the sensitivity of flows to recent performance.

This analysis is important because it sheds light on investors' short-termism and its associated negative consequences such as myopic investment decisions taken by managers.[48] If investors indeed use short-term results as a signal for investment decisions, then fund managers will have incentives to perform in the short run. Consistent with this inference, Jin and Kogan (2007) demonstrates that funds with higher short term performance pressure focus on short term profits at the cost of long run profits.

My results suggest that investors' decisions are strongly governed by longer-term rather than shorter-term fund performance. More precisely, both the direction of future flows to the fund and the effect of shorter-term performance depend on a fund’s longer-term performance. In particular, long-term winner funds obtain positive net flows, even though they underperform in the recent quarter. Long-term losers experience outflows despite a superior performance in the recent quarter. But long-term winners gain more inflows if their recent performance is also good and long-term losers experience more outflows if their recent performance is also poor.

Chevalier and Ellison (1997) document that fund flows into young funds are more sensitive to recent performance than flows into older funds. In a recent paper, Huang, Wei, and Yan (2012) confirm this dampening effect of fund age on sensitivity of fund flows on recent performance. Combining size and age, Sawicki and Finn (2002) find that investors respond more strongly to recent performance of small and young funds than to recent performance of large and old funds. This finding supports the idea that investors attempt to infer the quality of a fund from past performance. Since the lack of long track records increases the uncertainty regarding fund quality, investors rely more on recent performance to form their expectations, and hence the sensitivity of fund flows to recent

[48] See, for example, Meulbroek et. al. (1990), Bushee (1998), and Gaspar, Massa, and Matos (2005).

performance increases. Different from these papers, I concentrate on the *content* of past track records rather than its existence, using the level of long-term performance. I find that favorable long-term performance actually reinforces the effect of recent performance on future fund flows, which is opposite of age effect. Furthermore, my additional test shows that the effect of long-term performance on the flow-recent performance relation do not change across different age or size groups, implying that investors do not ignore the long-term performance in younger funds.

In addition to large flows for recent winners, the literature documents an absence of net outflows in the face of recent poor performance. Several researchers give plausible explanations that contribute to this convexity in the flow-performance relation. Gruber (1996) suggests that some investors fail to respond to poor performance due to the influence of advertising, brokers' advice, or market frictions such as tax inefficiencies. Lynch and Musto (2003) argue that investors rationally choose not to respond to poor performance since they expect that funds will change their personnel or strategies that produce it. Alternatively, Goetzmann and Peles (1997) show that investors suffer from cognitive dissonance, where they become over-optimistic about fund returns to justify their poor decisions. Moreover**,** Hendricks, Patel, and Zeckhauser (1992) hypothesize that flows to mutual funds may reflect status quo bias, framing, and data packaging[49].

My analysis highlights the importance of long-term performance on the shape of the flow-performance relation. I show that the convexity is mainly driven by long-term winners who experience a short period of poor performance. Investors do not ignore the

[49] The framing/data packaging conjectures that people give their decisions based on the most direct evidence. Therefore, flows follow the performance rankings, which are available on a regular and timely basis. However, investors also tend to stick with the strategies already adopted because of a reluctance to depart from status quo.

poor recent performance. Instead, they put more weight on the positive long-term performance. This makes it difficult to properly interpret the relation between short-term performance and flows from a model that omits long-term performance.

One important consequence of the convexity in the flow-performance relation is a possible distortion in managerial incentives. Since investors reward high performance but do not punish poor performance, fund managers have a free option to increase the riskiness of the portfolio to enhance the expected flows. Brown, Harlow, and Starks (1996), and Chevalier and Ellison (1997) find evidence of this excessive risk taking behavior in the funds with poor performance. In a recent paper, Huang, Sialm, and Zhang (2011) investigate the performance consequences of such risk-shifting. They document that funds that increase risk perform worse than those that keep risk stable over time. My results imply that greater risk-taking is not a completely free option for all funds. For instance, if a recent loser fund deliberately increases its risk and underperforms in the future as well, then the fund loses assets as investors eventually withdraw their capital.

Prior research has also investigated investors' sophistication using the performance of flows as a metric. Gruber (1996) and Zheng (1999) find evidence that the risk-adjusted short term performance of funds that experience inflows is better than that of funds that experience outflows. On the other hand, Sapp and Tiwari (2004) argue that this effect is due to a return momentum of the stocks in those funds. Additionally, Frazinni and Lamont (2008) document that the smart money is, at best, confined to very short time periods. Even though there is empirical controversy in the performance of fund flows, it raises the possibility that fund flows may cause fund performance rather than performance causes fund flows. Thus, the causal relation between performance and fund flows is a key point in the interpretation of flow-performance relation, yet this seems to be largely ignored by previous research. To partially address this concern, I conduct Granger causality tests in a VAR framework. Using fund portfolios sorted on size, age and styles, I find that performance Granger causes future fund flows, but there is no evidence of a reverse causality.

Overall, my findings are consistent with the idea that investors use information in past returns as signals about manager skills and future returns, as previously suggested in the literature (Berk and Green (2004), Gruber (1996), Sirri and Tufano (1998), Chevalier and Ellison (1997)). Complementing those studies, I find that the relationship between past returns and future flows is driven mainly by longer-term rather than shorter-term performance. This suggests that short-term oriented fund managers will find it more difficult to attract assets than long-term oriented fund managers.

The remainder of the essay is organized as follow: Section II outlines the data and the variables used in the analysis. Section III is devoted to the results and section IV summarizes the findings.

Data

I obtain the mutual fund data from the CRSP Survivorship Bias Free Mutual Fund Database, which includes information on fund returns, total net assets (TNA), fund fees, investment objectives, and other fund characteristics. The sample spans the period 1993-2010, during which CRSP provides consistent classifications of fund investment objectives. To compare the results with previous studies, I focus on domestic open-end equity funds. Therefore, sector funds, balanced funds, international funds, and bond funds are excluded from the sample. In constructing the sample, I follow Pastor and Stambaugh (2002) as closely as possible. Details about data selection are provided in the appendix.

CRSP treats different share classes of the same fund as distinct funds. Since the classes corresponding to a single fund share the same asset portfolio, their returns are correlated. But different share classes may cater to investors with different wealth, different investment purposes, and different levels of sophistication (Nanda, Wang, and Zheng (2009)). To capture these potential differences across investors, I conduct the analysis at the share-class level. Nevertheless, all results are qualitatively similar if I combine the multiple share classes of each fund.

I define the monthly flow as the net growth rate of total net assets (TNA):

$$\mathrm{Flow}_{i,t} = \frac{\mathrm{TNA}_{i,t} - \mathrm{TNA}_{i,t-1} * (1 + \mathrm{R}_{i,t})}{\mathrm{TNA}_{i,t-1}} \tag{2.1}$$

where $\mathrm{R}_{i,t}$ is the monthly return of fund i during month and $\mathrm{TNA}_{i,t}$ is the total net asset value of fund i at the end of month t This definition adjusts the total net asset for its appreciation during month t and therefore reflects the growth rate of the fund due to new investments. By using this formula, I implicitly assume that all new investments arrive at the end of each quarter. Following Huang, Wei, and Yan (2007), and Spiegel and Zhang (2010) I filter out the top and bottom of 2.5% tails of the flow data in order to mitigate the impact of potential outliers around mergers.[50]

I use two monthly performance measures: (i) raw return, (ii) risk-adjusted returns. To compute a month's risk-adjusted return, I first regress excess returns during the preceding 36 months (i.e. from t-36 to t-1) onto Carhart's (1997) four factors -market,

[50] Alternatively, I filter out the top and bottom of 1% tails of the flow data. Moreover, rather than drop outlying flow observations, I conduct the analysis using flow data that are winsorized at these levels. In all cases, results are similar.

size, value, and momentum- and obtain the estimates of factor loadings. To minimize estimation errors in the factor loadings, I require at least 24 months of valid returns in the estimation period. Then, I estimate the risk-adjusted return in each month t as the difference between the fund's realized excess return and the required return, defined as the factor loadings times the corresponding factor realizations. Size is defined as the total net assets (TNA) of a fund. Age is the number of years since the inception date of a fund.[51] Expense ratio is the fund total operating expenses expressed as a percentage of the fund's average net assets. Front load (rear load) measured as a percentage of fund assets is the fee investors pay when they buy (redeem) fund shares. Return is the monthly return of a fund net of expenses.

Fund characteristics are summarized in Table 2.1. Panel A reports cross-sectional mean and median values of various fund characteristics while Panel B shows correlations among the same variables. The number of funds has increased substantially from 731 in 1993 to 4295 in 2010. In all years, the median of total net assets per fund is much smaller than its corresponding mean, strongly suggesting that assets in the mutual fund industry `are concentrated in some big funds. The average age of the funds decreases until 2004 and increases afterwards. Fund fees including expense ratios and load fees have relatively remained stable over time.

The correlations between fund characteristics reported in Panel B are similar to those in Kacperczyk, Sialm, and Zheng (2005) and Chen et. al. (2004). In particular, fund flows are negatively correlated to fund size, age, turnover, and load fees. However, there

---

[51] I use "*first_offer_date*" as a proxy for the fund inception date.

is a positive correlation between expense ratios and fund flows. [52] All but one of the correlations are statistically significant at the 5% level, suggesting that these characteristics should be controlled for in multivariate tests.

Empirical Tests

Double Sorts

In this section, I analyze how future fund flows depend on past returns. I construct portfolios sorted on recent and long-term performance. Each month t, I compute recent performance as the average raw or risk-adjusted return of a fund over the past three months (from t to t-2). Long-term performance is the average of raw or risk-adjusted returns of past 24 months from t-27 to t-3.[53] To construct the portfolios I proceed as follows: In each month, I rank all funds into quintiles based on their long-term performance. Then, within each quintile, I further divide funds into quintiles according to their most recent performance. This creates 25 portfolios, which are held for one month after the portfolio formation. Table 2.2 reports the average monthly flows for these portfolios in the subsequent month. LT1 (LT5) refers to worst (best) group according to its long-term performance averages, while ST1 (ST5) is the worst (best) group based on its most recent performance. In Panel A1, performance is measured by raw returns, while in Panel A2 performance is measured by risk-adjusted returns. In addition to sequential

[52] The positive correlation between the expense ratio and the fund flow is unexpected. Chen et. al. (2004) also document a positive correlation between these two variables. However, these correlations show contemporaneous relationships between two variables. Since an increase in the expense ratio will decrease the net return in the same period, it may cause, all other factors being equal, a mechanical increase in our flow calculation due to the second term in our flow formula.

[53] I compute recent performance over the past quarter since fund managers are generally subject to quarterly relative performance monitoring. In a survey study, Baker (1998) reports this monitoring leads to more short-termist attitude and approach to the management of the funds. However, I repeat these tests using different horizons: one month and 12 months for recent performance, and 36 months for long-term performance. Results are similar when I use these different definitions and are available upon request.

Table 2.1. Summary Statistics

Panel-A presents the cross sectional mean (AVG) and median (ME) values of various fund characteristics at the end of each year from 1993/1 to 2010/12. N is the average number of funds. Size is the total net assets of funds (TNA). Age is the number of years since the inception date. Expense ratio is total annual management and administrative expenses. Front (rear) load is the fee that investors pay when they buy (redeem) fund shares. Return is the monthly return while flow is the monthly percentage change in TNA adjusted for investment returns. Turnover is defined as the minimum of purchases and sales over average TNA. All values except N, size and age are in percentages. Panel-B reports the contemporaneous correlations between these characteristics.

Panel A: Key statistics

Year	N	Size		Age		Expense Ratio		Turnover		Front Load		Rear Load		Return		Flow	
		AVG	ME	AVG	ME	AVG	ME	AVG	ME	AVG	ME	AVG	ME	AVG	ME	AVG	ME
1993	808	601.98	151.77	15.01	8.67	1.20	1.15	77.60	57.10	1.80	0.00	0.56	0.00	2.73	2.50	0.31	-0.16
1996	1454	739.73	143.19	11.01	5.08	1.34	1.24	86.32	65.63	1.37	0.00	0.72	0.00	-0.7	-0.9	0.45	-0.16
1998	2246	836.00	124.40	9.36	5.25	1.35	1.25	86.45	69.00	1.11	0.00	0.86	0.00	6.84	6.28	-0.3	-0.83
2000	3205	739.85	99.40	8.68	5.37	1.40	1.31	109.7	79.00	1.08	0.00	0.99	0.00	4.56	4.31	0.78	-0.18
2002	3944	417.52	59.20	8.79	6.05	1.48	1.38	105.9	75.00	1.12	0.00	1.06	0.00	-5.4	-5.4	0.21	-0.49
2004	4754	567.82	77.70	9.00	6.96	1.46	1.37	84.91	63.00	1.07	0.00	1.12	0.00	3.43	3.43	0.34	-0.55
2006	5081	657.39	82.70	9.70	7.75	1.37	1.30	81.62	62.00	1.03	0.00	1.05	0.00	0.69	0.72	-0.38	-0.79
2008	5171	404.33	57.30	10.40	8.25	1.25	1.20	87.12	63.00	0.98	0.00	0.79	0.00	3.21	2.99	-0.02	-0.84
2010	5427	537.23	72.60	11.15	9.27	1.24	1.20	79.22	58.00	0.92	0.00	0.62	0.00	6.30	6.45	0.40	-0.46

Panel B: Piecewise Correlations

	Size	Age	Expense Ratio	Turnover	Front Load	Rear Load	Return	Flow
Size	1.000							
Age	0.392	1.000						
Expense Ratio	-0.280	-0.325	1.000					
Turnover	-0.119	0.006	0.111	1.000				
Front Load	0.114	0.238	-0.518	0.037	1.000			
Rear Load	0.053	0.069	-0.187	-0.032	0.458	1.000		
Return	0.009	0.012	0.007	-0.014	-0.006	0.007	1.000	
Flow	-0.011	-0.096	0.025	-0.035	-0.022	-0.003	0.079	1.000

sorts, I also construct 25 portfolios using independent sorts on recent and long-term performance. Panel B1 and B2 of Table 2.2 reports the average monthly flows for these portfolios In Panel B1, performance is measured by raw returns, while in Panel B2 performance is measured by risk-adjusted returns.

The results suggest that long-term winners are punished for poor recent performance, because they receive fewer inflows in the next period than long-term winners with good recent performance. Similarly, long-term losers are punished more if they also have poor recent performance. But the intriguing result is that long-term winners receive inflows and long-term losers experience outflows, regardless of their recent performance. For example, funds in the top long-term quintile based on risk-adjusted returns (Panel A2) have positive net flows varying between 1.28% (for the worst recent performance) and 4.13% (for the best recent performance). In contrast, funds in the worst long-term quintile experience outflows between 1.29% and 0.05%, depending on their recent performance. Results are similar when I use independent sorts (see Panel-B2).

Changing perspectives, these results also show that short-term returns are more important for the best long-term performers. Specifically, the positive relation between recent winners and future flows can mostly be attributed to flows into funds that are also long-term winners. In particular, 60% of total flows in ST5 in Panel A2 come from the funds which are also in LT5, while the recent winners with average long-term performance (LT3) attract 13% of total flows in ST5. Moreover, the flow differential between ST1 and ST5 increases monotonically from 1.24% to 2.84% as long-term performance increases. The results in Panel A1 are similar. Hence, a superior return after

Table 2.2. Average Monthly Flow of Funds Sorted on Recent and Long-term Performance Rankings

This table presents the average monthly flow ratios of funds sorted by recent performance and long-term performance rankings. The sample period is from 1993/1 to 2010/12. Each month t, recent performance is taken as the average raw or risk-adjusted returns of a fund over the past three months (from t to t-2). Long-term performance is the average of raw or risk-adjusted returns of past 24 months from t-27 to t-3. To construct the portfolios, in each month, I first rank all funds into quintiles based on their long-term performance. Then, within each quintile, I further divide funds into quintiles according to their most recent performance. This creates 25 portfolios, which are held for one month after the portfolio formation. Alternatively, I construct 25 portfolios using independent sorts. In Panel A1 and B1, performance rankings are based on raw returns. In Panel A2 and B2 performance rankings are based on risk-adjusted returns. LT1 (LT5) refers to worst (best) group according to its long-term performance averages, while ST1 (ST10) is the worst (best) group based on its most recent performance ST5-ST1 presents the difference of flow ratios between ST5 and ST1; LT5-LT1 presents the difference of flow ratios between LT5 and LT1.Newey and West (1987) adjusted t-statistics (up to 3 lags) are reported in parenthesis. All variables are in percentage terms.

Panel A1: Dependent Sorts (Raw returns)

	LT1	LT2	LT3	LT4	LT5	LT5-LT1	
ST1	-1.25	-0.81	-0.26	0.35	1.27	2.52	***(11.3)***
ST2	-1.06	-0.31	0.20	0.90	1.88	2.94	***(14.3)***
ST3	-0.88	-0.15	0.56	1.05	2.31	3.19	***(14.1)***
ST4	-0.60	0.15	0.74	1.33	2.86	3.46	***(13.7)***
ST5	-0.17	0.53	1.25	2.03	4.81	4.98	***(17.9)***
ST5-ST1	1.08	1.34	1.51	1.68	3.54		
	(6.4)	***(9.8)***	***(10.3)***	***(11.0)***	***(15.3)***		

Panel A2: Dependent Sorts (Risk-adjusted returns)

	LT1	LT2	LT3	LT4	LT5	LT5-LT1	
ST1	-1.29	-0.72	-0.21	0.27	1.28	2.57	***(19.6)***
ST2	-0.74	-0.38	0.24	0.63	1.96	2.70	***(14.8)***
ST3	-0.51	-0.02	0.25	0.90	2.33	2.84	***(22.9)***
ST4	-0.44	-0.03	0.51	1.08	2.77	3.21	***(18.4)***
ST5	-0.05	0.26	0.85	1.61	4.13	4.18	***(18.4)***
ST5-ST1	1.24	0.98	1.06	1.34	2.85		
	(10.5)	***(10.8)***	***(10.9)***	***(13.5)***	***(15.1)***		

Panel B1: Independent Sorts (Raw returns)

	LT1	LT2	LT3	LT4	LT5	LT5-LT1	
ST1	-1.28	-0.84	-0.38	0.23	1.08	2.35	***(13.4)***
ST2	-1.08	-0.37	0.19	0.84	2.06	3.13	***(15.2)***
ST3	-0.81	-0.08	0.57	1.00	2.26	3.07	***(21.8)***
ST4	-0.62	0.19	0.76	1.48	2.74	3.36	***(24.3)***
ST5	-0.02	0.59	1.38	2.09	4.73	4.75	***(21.9)***
ST5-ST1	1.26	1.44	1.75	1.86	3.65		
	(7.0)	***(6.8)***	***(7.8)***	***(10.3)***	***(14.3)***		

Panel B2: Independent Sorts (Risk-adjusted returns)

	LT1	LT2	LT3	LT4	LT5	LT5-LT1	
ST1	-1.19	-0.73	-0.24	0.28	1.27	2.46	***(19.9)***
ST2	-0.55	-0.33	0.18	.58	1.96	2.52	***(18.1)***
ST3	-0.54	-0.08	0.28	0.85	2.28	2.83	***(20.5)***
ST4	-0.40	-0.04	0.52	1.08	2.58	2.98	***(23.8)***
ST5	-0.01	0.35	0.91	1.66	3.78	3.78	***(23.3)***
ST5-ST1	1.19	1.08	1.16	1.38	2.51		
	(9.4)	***(11.3)***	***(10.8)***	***(11.3)***	***(14.5)***		

a period of a good performance has more influence on flows than a superior return after a period of mediocre performance. This is consistent with the idea that investors divert their capital to the funds in which they are confident about the manager's ability.

Regression Analysis

In this section, I adopt a regression approach to control for multiple fund characteristics simultaneously. I first analyze the impact of long-term performance on the overall level of sensitivity of flows to recent performance. In particular, I run a regression that includes recent performance, long-term performance and the interaction terms between these variables. Later in the paper, I consider a piecewise-linear specification used in Sirri and Tufano (1998). Specifically, recent performance is divided into three regions- Low, Mid, High- to capture the documented convex flow-performance relation. I use these three regions, interaction terms between these regions and long-term performance, and long-term performance itself. This specification tests how the long-term performance affects the convexity of flow-performance relation.

The Effect of Long-term Performance on the Overall Flow-Performance Relation

To estimate the effect of long-term performance on the relation between flows and short term performance, I run the following specification:

$$\begin{aligned}\text{Flow}_{i,t} = \alpha + \beta_1 \times \text{STPerf}_{i,t} + \beta_2 \times \text{LTPerf}_{i,t} + \beta_3 \times \text{STPerf}_{i,t} \times \text{LTPerf}_{i,t} \\ + \text{Controls}_{i,t}\end{aligned} \tag{2.2}$$

where $\text{STPerf}_{i,t}$ and $\text{LTPerf}_{i,t}$ are short-term and long-term performance rankings of fund i in month t. To compute short-term performance rankings ($\text{STPerf}_{i,t}$), I proceed as follows: Each month t, I rank all funds according to their recent performance, measured as the average raw or risk-adjusted returns during previous three months (from t-2 to t).

Then I assign performance ranks ($\mathrm{STPerf}_{i,t}$), that are uniformly distributed between 0 (worst) and 1 (best). For raw returns, funds are ranked within their investment categories; for risk adjusted, all funds are ranked together. Similarly, $\mathrm{LTPerf}_{i,t}$ is the performance rank of funds (from 0 to 1) according to the raw or risk-adjusted returns during the previous 24 months from t-27 to t-3. The coefficient of the interaction term- β_3 - captures the conditional effect of long-term performance on the sensitivity of flows to recent performance. Control variables include the total risk of a fund measured by the standard deviation of the returns over the preceding 24 months, logarithm of age, fund size measured by the logarithm of total net assets, expense ratio, load fees, monthly fund flows, and aggregate flow into each investment category. These variables are known to affect fund flows (Sirri and Tufano (1998), Huang, Wei, and Yan (2007), Del Guercio and Tkac (2008)). I also use dummy variables for investment styles (Bergstresser and Poterba (2002)). All control variables are lagged at least one month. To estimate the regression parameters, I use Fama and MacBeth's (1978) approach. Standard errors are corrected autocorrelation and heteroskedasticity using the method of Newey and West (1987) with three lags. In addition, I also estimate the regression using Pooled OLS with standard errors clustered by month.

Table 2.3 presents the parameter estimates as well as the corresponding t-statistics. The coefficients on control variables show that fund size negatively affect future flows. This is consistent with Berk and Green's (2004) model, in which flows cease due to the decreasing returns to scale in performance as size grows. Furthermore, consistent with the extant literature, older funds as well as funds with higher expenses and turnover and turnover have on average lower future flows.

Table 2.3. The Effect of Long-Term Performance on the Sensitivity of Fund Flows to Recent Performance

This table reports the parameter estimates from Fama and MacBeth's (1973) and pooled OLS regressions of monthly observations for funds from 1993/1 to 2010/12. The dependent variable is the monthly flow ratio. StPerf is the performance rank of funds (from 0 to 1) according to the raw or risk-adjusted returns during the previous three months from t-3 to t-1. LTPerf is the performance rank of funds (from 0 to 1) according to the raw or risk-adjusted returns during the previous 24 months from t-28 to t-4 Other control variables include volatility of the past 24-month returns, turnover, expense ratios, logarithm of fund age, and logarithm of fund size as proxied by funds total net assets. All control variables are lagged at least one month. All specifications include style dummies. Column (1) and (2) report parameter estimates from Fama and MacBeth's (1973) regressions. Newey and West (1987) adjusted t-statistics (up to 3 lags) are reported in parenthesis. Second column shows the coefficients from pooled OLS. Standard errors are clustered by month.

	Raw Returns		Risk-adjusted Returns	
	Fama-Macbeth	Pooled OLS	Fama-MacBeth	Pooled OLS
STPerf	0.006***	0.005***	0.005***	0.004***
	(5.77)	(5.29)	(8.84)	(7.66)
LTPerf	0.016***	0.013***	0.012***	0.011***
	(15.24)	(12.19)	(19.36)	(20.64)
STPerf*LTPerf	0.016***	0.014***	0.010***	0.010***
	(12.83)	(7.46)	(10.38)	(10.64)
Flow	0.402***	0.418***	0.414***	0.425***
	(53.91)	(62.40)	(52.36)	(63.38)
Log(Age)	-0.005***	-0.006***	-0.005***	-0.006***
	(-29.17)	(-32.81)	(-30.73)	(-32.80)
Log(Size)	-0.001***	-0.001***	-0.001***	-0.000***
	(-13.41)	(-8.10)	(-8.95)	(-4.95)
Cat. Flow	15.986	0.255***	15.978	0.197***
	(1.12)	(4.61)	(1.11)	(3.68)
Exp. Ratio	-0.312***	-0.470***	-0.327***	-0.454***
	(-5.83)	(-16.42)	(-6.82)	(-16.21)
Turnover	0.000	0.000	0.000	0.000*
	(0.64)	(0.62)	(0.93)	(1.88)
Volatility	-0.084***	0.007	0.037	0.010
	(-3.67)	(0.58)	(1.47)	(0.81)
Load	-0.008	-0.015***	-0.007	-0.014***
	(-1.38)	(-3.90)	(-1.25)	(-3.84)
N	659243	659243	636179	636179
R-sq	0.29	0.28	0.28	0.27

Regarding the flow-performance relationship, the coefficients of interaction terms - β_3 - are 0.016 in Fama and MacBeth's (1973) regression and 0.014 in pooled OLS, when raw returns are used. Both coefficients are statistically significant at the 1% level, suggesting that long-term performance increases the sensitivity of flows to recent performance. Results are similar when risk-adjusted returns are used. Moreover, the coefficients of long-term performance are three times greater than those of short-term performance. This implies that the effect of long-term performance on flows is greater than that of short-term performance. For instance, when raw returns are used, moving five percentiles in the long term rankings will increase flows by 0.12% per month for a fund that is in the 50th performance percentile according short-term rankings (≈5% x (0.016+0.016 x 0.5)). However, moving five percentiles in the short term rankings will increase flows 0.07% per month for a fund that is in the 50th percentile according to long-term performance (≈5% x (0.006+0.016 x 0.5)). Overall the results suggest that investors condition their response to recent performance on the long-term performance history.

Prior research shows that age affects the sensitivity of flow to recent performance (Chevalier and Ellison (1997, 1998), and Huang Wei, and Yan (2012). Specifically, fund flows into young funds are more sensitive to recent performance than flows into older funds. Huang, Wei, and Yan (2012) argue that age is related to the uncertainty associated with the fund quality, and since younger funds have less historical data, investors should rely on more recent performance in their decisions. In addition to age, Sawicki and Finn (2002) find that investors respond more strongly to recent performance of small and young funds than to recent performance of large and old funds. Age or size does not necessarily imply a better long-term performance. Nevertheless, in order to see whether investors capture the same information from long-term performance like from age or size, I divide all funds into age or size quintiles (Q1: lowest quantile, Q5: highest quintile) , and conduct my tests within each quintile. Table 2.4 shows the parameter estimates of

STPerf, LTPerf, and STPerf x LTPerf from Fama and Macbeth (1973) regressions. [54] Considering the age quintiles, we observe that the coefficients of STPerf decrease from 0.011 in Q1 to 0.005 in Q5 and the difference is statistically significant. The coefficient of interaction term does not change significantly even though there is a drop in magnitudes. This suggest that for the same long-term performance, the effect of recent performance on future flows is greater among young funds as documented in Chevalier and Ellision (1997,1998), and Huang, Wei, and Yan (2011). However, the effect of long-term performance on the flow-recent performance relation does not change across age quintiles. Similar results are obtained for the size quintiles. Overall, the findings suggest that the age or size effect does not subsume the effect of long-term performance on the sensitivity of flows to recent performance.

The Effect of Long-term Performance on the Convexity of Flow-Performance Relation

Previous literature documents that the sensitivity of future flows to recent past performance is not linear (Ippolito (1992), Chevalier and Ellison (1997), and Sirri and Tufano (1998)). To model the non-linearity in the flow-performance relation, I follow Sirri and Tufano (1998) and divide the short-term performance rankings $\mathrm{STPerf}_{i,t}$ into three groups: $\mathrm{Low}_{i,t} = \min(\mathrm{STPerf}_{i,t}, 0.2)$, $\mathrm{Mid}_{i,t} = \min(\mathrm{STPerf}_{i,t} - \mathrm{Low}_{i,t}, 0.6)$, $\mathrm{High}_{i,t} = \mathrm{Rank}_{i,t} - \mathrm{Mid}_{i,t} - \mathrm{Low}_{i,t}$. Using these fractional ranks, I run the following piecewise linear regression:

$$\begin{aligned}\mathrm{Flow}_{i,t} = \alpha + \beta_1 \times \mathrm{Low}_{i,t} + \beta_2 \times \mathrm{Mid}_{i,t} + \beta_3 \times \mathrm{High}_{i,t} + \beta_4 \times \mathrm{Low}_{i,t} \times \mathrm{LTPerf}_{i,t} \\ + \beta_5 \times \mathrm{Mid}_{i,t} \times \mathrm{LTPerf}_{i,t} + \beta_6 \times H\mathrm{igh}_{i,t} \times \mathrm{LTPerf}_{i,t} + \mathrm{Controls}_{i,t}\end{aligned} \tag{2.3}$$

[54] Pooled OLS results are similar and omitted for brevity.

Table 2.4. The Effect of Long-term Performance on the Sensitivity of Fund Flows to Recent Performance Across Age and Size Groups

This table reports the parameter estimates from Fama and MacBeth's (1973) for funds from 1993/1 to 2010/12 for different size and age quintiles. Each month, I sort the funds according to their total net assets or their age. Q1 refers to the group of funds with lowest total net assets (age), while Q5 is the group of funds with highest total net assets. Regressions are performed within each group. The dependent variable is the monthly flow ratio. StPerf is the performance rank of funds (from 0 to 1) according to the raw or risk-adjusted returns during the previous three months from t-3 to t-1. LTPerf is the performance rank of funds (from 0 to 1) according to the raw or risk-adjusted returns during the previous 24 months from t-28 to t-4 Other control variables (not reported) include volatility of the past 24-month returns, turnover, expense ratios, logarithm of fund age, and logarithm of fund size as proxied by funds total net assets. All control variables are lagged at least one month. All specifications include style dummies. Newey and West (1987) adjusted t-statistics (up to 3 lags) are reported in parenthesis. The table also reposts the the p-values from a Chow test, that determines whether the coefficients of estimates between Q5 and Q1 are statistically different from each other.

	Fama-Macbeth Regressions within each age quintile						Fama-Macbeth Regressions within each size quintile					
	Raw returns						Raw returns					
	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (p-values)	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (p-values)
STPerf	0.011***	0.007***	0.006***	0.004***	0.005***	**0.001**	0.008***	0.005***	0.004***	0.006***	0.005***	**0.016**
	(5.66)	(3.77)	(4.85)	(3.56)	(4.86)		(5.31)	(3.37)	(2.81)	(4.21)	(3.48)	
LTPerf	0.019***	0.019***	0.019***	0.015***	0.014***	**0.027**	0.017***	0.015***	0.014***	0.018***	0.015***	0.390
	(8.64)	(12.38)	(12.56)	(11.98)	(11.43)		(9.88)	(8.99)	(8.82)	(12.04)	(9.46)	
STPerf*LTPerf	0.016***	0.018***	0.017***	0.016***	0.012***	0.143	0.016***	0.021***	0.021***	0.016***	0.012***	0.139
	(5.53)	(8.11)	(8.90)	(8.23)	(6.90)		(6.67)	(7.61)	(9.40)	(7.77)	(6.54)	
	Risk-adjusted returns						Risk-adjusted returns					
	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (p-values)	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (p-values)
STPerf	0.008***	0.003***	0.005***	0.004***	0.005***	**0.037**	0.007***	0.004***	0.003***	0.004***	0.004***	**0.000**
	(6.48)	(3.40)	(4.73)	(4.69)	(6.88)		(5.84)	(4.00)	(3.11)	(5.77)	(3.99)	
LTPerf	0.010***	0.013***	0.015***	0.014***	0.012***	0.081	0.012***	0.011***	0.011***	0.013***	0.012***	0.438
	(7.55)	(11.23)	(11.68)	(12.19)	(13.07)		(8.78)	(8.85)	(10.02)	(13.11)	(12.23)	
STPerf*LTPerf	0.009***	0.014***	0.011***	0.010***	0.006***	0.275	0.009***	0.013***	0.012***	0.010***	0.008***	0.980
	(4.35)	(7.13)	(6.13)	(4.85)	(4.52)		(4.00)	(6.48)	(5.23)	(6.34)	(5.50)	

where $\text{LTPerf}_{i,t}$ is the performance rank of funds (from 0 to 1) according to the raw or risk-adjusted returns during the previous 24 months from t-27 to t-3. This specification allows the sensitivity of flow to recent performance to change at three different performance levels. Hence, we can analyze whether long-term performance has different effects among short-term winners and losers.

To confirm the convexity of flow-performance relation in the sample, I first run regressions without interaction terms. Table 2.5 presents the parameter estimates as well as the corresponding t-statistics for Fama and MacBeth's (1973) and pooled OLS regressions. In all specifications, the flow sensitivity to different performance levels is positive and significant. However, it is much higher for High than for Mid and Low. For instance, when raw returns are used, the coefficient of High is 0.055 in the Fama and Macbeth's (1973) regression, while it is 0.049 in the pooled regression. The corresponding coefficients for Low are 0.022 and 0.014. Results for risk-adjusted returns are very similar. A Wald test of the equality of coefficients between High and Low is rejected at 1% significance level.[55]

While all performance categories experience inflows, I also find that low performance leads to greater inflows than average performance. This differs from Sirri and Tufano (1998), who find an insignificant coefficient for the Low rank. In this respect, my results are very similar to those in Huang, Wei, and Yan (2007) and Spiegel and Zhang (2010), who find a positive significant coefficient for the Low rank as well. Actually, Huang, Wei, and Yan (2007) argue that in the 1990s, flows become significantly more sensitive in the low and medium performance ranges due to the

[55] The sensitivity of Mid rank is lower than Low rank, suggesting concavity as performance decreases. This pattern is also documented in Huang, Wei, and Yan (2007), and Spiegel and Zhang (2010). However, my focus is mainly in the relation between High and Low.

Table 2.5. Unconditional Flow-Performance Relation

This table reports the parameter estimates from Fama and MacBeth's (1973) and pooled OLS regressions of monthly observations for funds from 1993/1 to 2010/12. The dependent variable is the monthly flow ratio. Low, Mid, and High correspond to fractional rankings of raw returns or risk-adjusted returns over the previous three months. Other control variables include volatility of the past 24-month returns, turnover, expense ratios, logarithm of fund age, and logarithm of fund size as proxied by funds total net assets. All control variables are lagged one month. All specifications include style dummies. Column (1) and (2) report parameter estimates from Fama and MacBeth's (1973) regressions. Newey and West (1987) adjusted t-statistics (up to 3 lags) are reported in parenthesis. Second column shows the coefficients from pooled OLS. Standard errors are clustered by month.

	Raw Returns		Risk-adjusted Returns	
	Fama-Macbeth	Pooled OLS	Fama-MacBeth	Pooled OLS
Low	0.022***	0.014***	0.021***	0.019***
	(7.58)	(7.07)	(10.36)	(11.59)
Mid	0.011***	0.008***	0.005***	0.005***
	(11.29)	(13.11)	(12.14)	(12.12)
High	0.055***	0.049***	0.044***	0.041***
	(18.92)	(16.61)	(13.80)	(15.64)
Flow	0.432***	0.443***	0.438***	0.448***
	(57.39)	(63.84)	(54.86)	(61.91)
Log(Age)	-0.006***	-0.006***	-0.005***	-0.006***
	(-31.62)	(-36.04)	(-29.71)	(-34.50)
Log(Size)	-0.000***	-0.000**	-0.000***	-0.000
	(-5.18)	(-2.08)	(-4.02)	(-1.60)
Car. Flow	15.643	0.214***	15.597	0.203***
	(1.08)	(4.05)	(1.13)	(3.81)
Exp. Ratio	-0.370***	-0.497***	-0.378***	-0.503***
	(-7.84)	(-17.95)	(-8.19)	(-18.09)
Turnover	-0.000***	-0.000	-0.000**	-0.000
	(-2.96)	(-1.36)	(-2.23)	(-1.30)
Volatility	-0.020	0.001	0.013	0.006
	(-0.78)	(0.09)	(0.48)	(0.46)
Load	-0.010*	-0.014***	-0.011*	-0.017***
	(-1.68)	(-3.74)	(-1.90)	(-4.52)
N	662205	662205	648140	648140
R-sq	0.28	0.27	0.27	0.26

substantial decrease in participation costs. This supports the idea that investors may have certain level of confidence about managers' ability. Hence, some investors that have a certain level of sophistication may choose to incorporate long-term performance history into their evaluation of the quality of recent signals and therefore, into investment decision process.

Next, I run the conditional regressions that include the interaction terms. Table 2.6 reports the parameter estimates and the corresponding t-statistics. The interaction terms are positive and significant, meaning that as long-term performance increases, the sensitivity of flows to recent performance also increases in all performance levels. Using parameter estimates in column (1), the sensitivity of fund flows to recent performance changes in the lowest performance category Low is given by:

$\frac{\partial \text{Flow}}{\partial \text{Low}} = 0.010 + 0.021 * \text{LTPerf}$. This means any performance increase in the lowest performance category is associated with significantly greater inflows if the fund has a good long-term performance than a similar move if the fund has a bad long-term performance. For instance, consider that a fund in the Low category moves from 10^{th} percentile to 15^{th} percentile.[56] This five percentiles move will increase fund flows by 0.15% per month ($\approx$5% x (0.010 + 0.021 x 0.9)) if the fund is in the 90^{th} percentile according to its long-term performance. The corresponding increase in fund flows is 0.06% per month if the fund is in the 10^{th} percentile according to its long-term performance. The impact of long-term performance is even greater in the High category. The sensitivity of fund flows to performance change in this category is given by:

$\frac{\partial \text{Flow}}{\partial \text{High}} = 0.011 + 0.076 \times \text{LTPerf}$. Hence, long-term performance affects the relation between recent performance and flows more among funds with better recent performance. By the same token, given two funds that are recent losers (winners), the one

[56] Note that Low can take values from 0 to 0.2 (20^{th} percentile).

Table 2.6. Conditional Flow-Performance Relation

This table reports the parameter estimates from Fama and Macbeth (1973) and pooled OLS regressions of monthly observations for funds from 1993/1 to 2010/12. The dependent variable is the monthly flow ratio. Low, Mid, and High correspond to fractional rankings of raw returns or risk-adjusted returns over the previous three months. Other control variables include volatility of the past 24-month returns, turnover, expense ratios, logarithm of fund age, and logarithm of fund size as proxied by funds total net assets. All control variables are lagged one month. All specifications include style dummies. Column (1) and (2) report parameter estimates from Fama and MacBeth's (1973) regressions. Newey and West (1987) adjusted t-statistics (up to 3 lags) are reported in parenthesis. Second column shows the coefficients from pooled OLS. Standard errors are clustered by time.

	Raw Returns		Risk-adjusted Returns	
	Fama-Macbeth	Pooled OLS	Fama-MacBeth	Pooled OLS
Low	0.010***	0.013***	0.016***	0.015***
	(2.73)	(3.75)	(4.96)	(5.62)
Mid	0.005***	0.004***	0.002**	0.001*
	(4.22)	(3.97)	(2.27)	(1.90)
High	0.011*	0.012**	0.021***	0.017***
	(1.82)	(2.14)	(4.86)	(4.66)
Low*LTPerf	0.021***	0.008	0.013**	0.011**
	(3.19)	(1.33)	(2.46)	(2.38)
Mid*LTPerf	0.007***	0.007***	0.006***	0.006***
	(3.69)	(3.49)	(4.47)	(4.33)
High*LTPerf	0.076***	0.060***	0.033***	0.036***
	(6.96)	(5.83)	(4.16)	(5.62)
LTPerf	0.017***	0.015***	0.012***	0.012***
	(13.84)	(10.43)	(13.32)	(14.46)
Flow	0.399***	0.417***	0.412***	0.424***
	(53.57)	(62.28)	(52.37)	(63.51)
Log(Age)	-0.005***	-0.006***	-0.005***	-0.006***
	(-29.28)	(-33.16)	(-30.29)	(-32.92)
Log(Size)	-0.001***	-0.001***	-0.001***	-0.000***
	(-13.13)	(-8.09)	(-8.83)	(-5.01)
Car. Flow	14.788	0.257***	15.229	0.197***
	(1.08)	(4.62)	(1.10)	(3.67)
Exp. Ratio	-0.322***	-0.482***	-0.335***	-0.465***
	(-5.94)	(-16.93)	(-6.92)	(-16.75)
Turnover	0.000	0.000	0.000	0.000
	(0.32)	(0.21)	(0.75)	(1.63)
Volatility	-0.099***	0.004	0.028	0.008
	(-4.14)	(0.30)	(1.11)	(0.61)
Load	-0.007	-0.014***	-0.007	-0.014***
	(-1.18)	(-3.64)	(-1.23)	(-3.72)
N	659243	659243	636179	636179
R-sq	0.29	0.28	0.28	0.27

with a worse long-term performance will lose more flows than the one with a better performance history.

To better understand the conditional effect of long-term performance, I plot the sensitivity of fund flows to recent performance as a function of long-term performance. Specifically, I gradually increase long-term performance ranking (LTPerf) from 0 to 1 by a factor of 0.1 and, for each LTPerf percentile, I compute the sensitivity of flows in in the Low and Hig category. [57] Figure 2.1 plots the sensitivity of flows in the High and Low as well as the difference (convexity) between the sensitivity of flows in these two categories using parameter estimates from Fama and MacBeth's (1978) regressions. Panel A uses rankings computed from raw returns, while Panel B uses rankings computed from risk-adjusted returns. Both graphs suggest that flows' response to performance gets stronger in High and Low as long term rankings increase. However, the increase in High category is greater than the increase in the Low category. This suggests that long-term performance affects the convexity of flow performance relation.

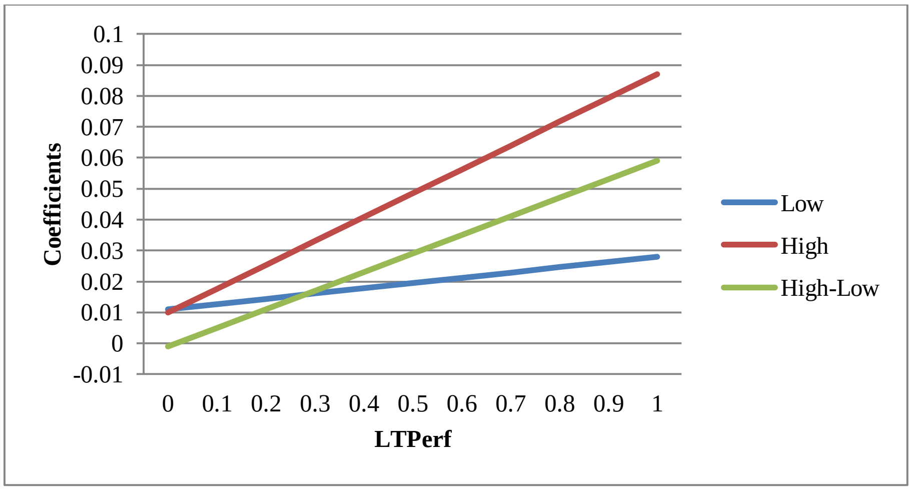
In my first essay, I show that maximum-style adjusted daily returns over a month affect future fund flows. This finding is in line with the idea that investors have a preference for lottery-like payoffs, which are high positive payoffs with relatively small probabilities. In order to see whether my results are robust to control for lottery-like preferences, I run my Fama and Macbeth’s (1973) regressions after including maximum style-adjusted daily returns (MAX) over a month as a control variable. Table 2.7 shows the coefficients of MAX, long-term performance and interaction of long-term performance with recent performance.

[57] For instance, when LTPerf is 0.2, then the total effect of Low is 0.011+0.017*0.2=0.014 based on the parameter estimates in column (1) of Table 2.6.

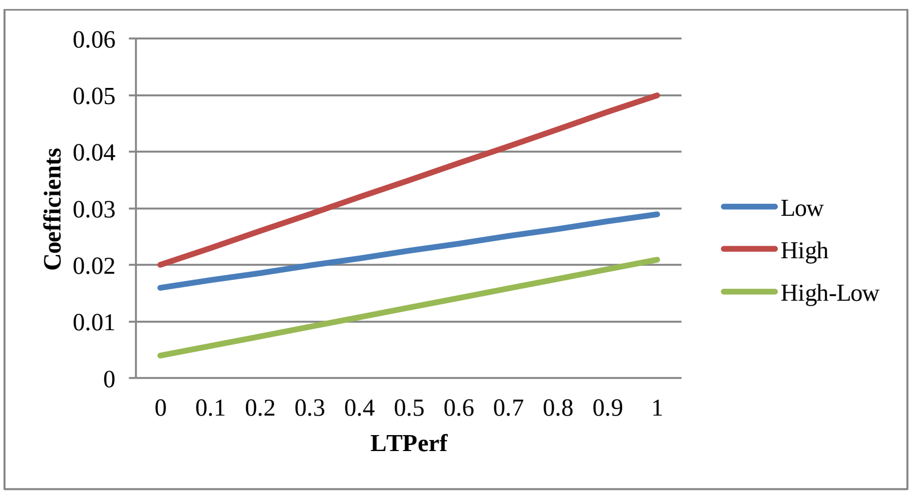
Figure 2.1. Sensitivity of Fund Flows to Recent Performance as a Function of Long-term Performance

This figure plots the sensitivity of flows to recent performance in the High and Low category as a function of long term performance ranking (LTPerf). Long term rankings are gradually increased from 0 to 1 by a factor of 0.1. Then, for each percentile of long term rankings, the total effect of recent performance onto flows is calculated using the coefficient estimates from Fama and MacBeth's (1973) regression reported in column (1) and (2) of Table 2.6. Panel A use coefficients estimated when raw returns are used; Panel B use coefficients estimated when risk-adjusted returns are used. In both panels, High-Low refers to the difference (convexity) between the sensitivity of flows in these two categories.

Panel A: Rankings based on raw returns



Panel B: Rankings based on risk-adjusted returns



The results show that MAX and long-term performance as well as the interaction terms are highly significant, suggesting that no variable subsume the effect of other variable. Presumably, consumers gather information from a variety of sources about alternative mutual funds and develop a set of fund attributes that are important to them from their information set. In this process, they may well adopt a multi-attribute model in determining their choices (Fishbein and Azjen (1975), Lanchaster (1966), and Capon, Fitzsimons, and Prince (1996)). As a result, the existence of a preference for lottery-like returns does not necessarily preclude that some investors would positively value long-term performance or vice versa. The value that investors attach to long-term performance and lottery-like returns may differ among different clientele. Accordingly, Genc (2012) show that preference for lottery-like payoffs only exist in retail funds, which predominantly cater to individual investors. Nevertheless, as long as long a sizable portion of investors consider these fund characteristics in their purchase decisions, we can observe the effect of both fund characteristics on fund flows. Moreover, it is also important to emphasize that since investors' utility functions are unobservable, one cannot argue a preference for lottery-like payoffs show investors' unsophistication. As discussed in my first essay, investors may derive utility from this preference and they optimally choose these funds as a result of a utility maximization. Actually, it is also arguable that following long-term performance might be a non-optimal decision given the fact that long-term performance in mutual funds does not persist (Carhart (1997)).

Granger Causality Tests

My multivariate regressions reflect correlations between performance and future fund flows, but they do not reveal direct information about causality between these variables.

Table 2.7. The Effect of Long-term Performance after Controlling for MAX

This table reports the parameter estimates from Fama and Macbeth (1973) regressions of monthly observations for funds from 1998/9 to 2010/12. StPerf is the performance rank of funds (from 0 to 1) according to the raw or risk-adjusted returns during the previous three months from t-3 to t-1. LTPerf is the performance rank of funds (from 0 to 1) according to the raw or risk-adjusted returns during the previous 24 months from t-28 to t-4.Low, Mid, and High correspond to fractional rankings of raw returns or risk-adjusted returns over the previous three months. Other control variables (omitted for brevity) include volatility of the past 24-month returns, turnover, expense ratios, logarithm of fund age, and logarithm of fund size as proxied by funds total net assets. All control variables are lagged one month. All specifications include style dummies. Newey and West (1987) adjusted t-statistics (up to 3 lags) are reported in parenthesis.

	Raw returns		Risk-adjusted returns	
MAX	0.276***	0.229***	0.260***	0.235***
	(7.31)	(6.34)	(7.30)	(6.92)
STPerf	0.005***		0.004***	
	(4.10)		(6.73)	
LTPerf	0.015***		0.011***	
	(11.63)		(16.16)	
STPerf*LTPerf	0.015***		0.009***	
	(11.47)		(8.26)	
Low		0.012***		0.012***
		(3.10)		(4.41)
Mid		0.005***		0.002*
		(3.09)		(1.82)
High		0.013**		0.020***
		(2.25)		(4.60)
Low*LTPerf		0.015**		0.015***
		(2.14)		(3.47)
Mid*LTPerf		0.008***		0.006***
		(4.09)		(3.91)
High*LTPerf		0.062***		0.028***
		(6.69)		(3.09)

For instance, Gruber (1996) shows that funds with higher cash flows perform better in the future, indicating that fund flows can be used as the predictor of fund performance. Even though the "smart money" effect of Gruber (1996) is shown to be short-lived, it is confined to small funds and partially explained by strategies of betting on winners (Zheng (1999), Sapp and Tiwari (2004), Frazzini and Lamont (2008)). This raises the possibility of causality going in the opposite direction i.e. that flows lead to superior performance at least in the short run. Therefore, in this section, I look at this issue by performing Granger causality tests.

Granger causality test determines whether one time series is useful in forecasting another. For example, a time series X is said to Granger cause another time-series Y if lagged values of X (in the presence of lagged values of Y) provide information about the future values of Y. To conduct the test, I first obtain a time-series of performance and flows for fund portfolios sorted on (a) size (b) age, and (c) styles. Chevalier and Ellison (1997) study the flow-performance relation across age regions. They show that flows to young funds are more sensitive to recent performance. Combining size and age, Sawicki and Finn (2002) find that investors respond more strongly to recent performance of small and young funds than to recent performance of large and old funds. To control for this size and age effect on the flow-performance relation, I first sort the fund into quintiles based on size or age (i.e. five size and five age portfolios). I also create fund portfolios based on five styles- capital appreciation, growth and income, income, growth, and small growth. For each of these fifteen portfolios, I calculate the performance using the average of three month returns and average fund flows. Then, I run Granger causality test using a Vector Auto Regression (VAR) framework. The test examines the causal relation between the average of prior three month return and fund flows, and between the fund flows and the average of the following three month returns. To estimate VAR properly, stationary data is needed. Augmented Dickey Fuller tests show no evidence of non-stationarity in both return and flow series.[58] Another practical issue in VAR is

[58] In all size and age quintiles, and style groups, the null hypothesis that return (flow) series has a unit root is rejected at 1% level. Only exception is the null hypothesis is rejected at 5% for the flow series in the lowest two size quintiles and the highest age quintiles.

determining the number of lags used for the endogenous variables. Too many lags could increase the errors in the forecasts; too few could leave out relevant information. I use Akaike's information criterion to select the number of lags.[59] Table 2.8 reports the χ^2 and the corresponding p-values. In all tests, the null hypothesis, that prior three month performance does not Granger cause flows (Perf $\nrightarrow$ Flow), is rejected at 1%, while the hypothesis that flow does not Granger cause performance cannot be rejected. Thus, the table shows a strong one way Granger causality from performance to flows, while there is no evidence of a reverse causality.

Chapter Conclusion

Previous research shows that investors are insensitive to recent poor performance but disproportionally flock to funds with the best recent performance. In this paper, I study how a fund's longer-term performance affects this relationship. I find that investors take into account recent performance, but the main driver of future fund flows is longer-term historical performance. Specifically, long-term winners can afford a losing quarter, while long-term losers cannot increase their flows by boosting their recent performance. These findings imply that short-termism on behalf of fund managers may not be a competitively viable strategy, because fund investors see through attempts to boost short-term performance at the expense of long-term performance.

[59] Other commonly used information criteria are Schwarz's Bayesian information criterion (SBIC), and the Hannan and Quinn information criterion (HQIC). Using simulations, Ivanov and Kilian (2005) conclude that for monthly VAR models AIC tends to produce the most accurate estimates for realistic sample sizes. Nevertheless, the number of lags selected by these three methods generally coincides in my sample.

Table 2.8. Granger Causality Tests

This table shows Granger causality tests between short-term performance and percentage flows within (a) each size quintiles, (b) within each age quintiles (c) within each style. Q1 is the portfolio of funds with the smallest total net assets or age, Q5 is the portfolio of funds with greatest total net assets or age. Short-term performance is the average of three month returns and flow is the average of percentage flows.. In each case, I first test the null hypothesis that performance from t-3 to t-1 does not Granger cause flows in month t and then whether flow in month t does not Granger cause performance from t+1 to t+3. For each test, I report χ^2 and the corresponding p-value. I use Akaike's information criterion to determine the number of lags.

Granger Causality within each size quintiles

Null Hypothesis	Q1	Q2	Q3	Q4	Q5
Perf $\nrightarrow$ Flow					
χ^2	41.222	75.594	69.211	90.051	69.189
p-value	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Flow $\nrightarrow$ Perf					
χ^2	5.642	7.449	7.395	8.868	4.618
p-value	0.69	0.49	0.49	0.35	0.79

Granger Causality within age quintiles

Null Hypothesis	Q1	Q2	Q3	Q4	Q5
Perf $\nrightarrow$ Flow					
χ^2	89.545	78.972	49.538	48.48	36.685
p-value	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Flow $\nrightarrow$ Perf					
χ^2	9.682	7.573	5.464	5.789	5.905
p-value	0.28	0.48	0.71	0.67	0.66

Granger Causality within each styles

Null Hypothesis	CA	GI	GR	I	SG
Perf $\nrightarrow$ Flow					
χ^2	85.182	30.69	86.529	17.176	81.468
p-value	**0.00**	**0.00**	**0.00**	**0.03**	**0.00**
Flow $\nrightarrow$ Perf					
χ^2	10.9333	4.572	8.396	3.903	7.334
p-value	0.21	0.8	0.39	0.87	0.5

CHAPTER IV

CONCLUSION

In this dissertation I analyze how investors make their fund choices. In the first essay, I study whether investors' trading decisions are influenced by a preference for lottery-like features (i.e. high payoffs with small probabilities) in asset returns. Using maximum style-adjusted daily returns (MAX) as a proxy for lottery-like returns, I show that investors are attracted to funds that provide lottery-like returns. Specifically, I document a positive and significant relation between MAX and flows, which is robust to various control variables including fund performance, fund size, age, turnover, fund fees, (idiosyncratic) volatility, and (idiosyncratic) skewness of fund returns. This flow-MAX relationship exists only among funds that cater primarily to retail investors, suggesting that predominantly retail investors are subject to this type of preference. Moreover, I document that MAX in a period predicts MAX in future periods. Therefore, if lottery-like features enter investors' utility function and valuable to investors, then choosing such funds may be optimal. However, this MAX-based strategies are costly to investors as funds with high MAX subsequently underperform otherwise similar funds on a risk-adjusted basis.

In the second essay, I investigate the role of recent and long-term performance on future fund flows using a conditional setting. Previous research shows that investors are insensitive to recent poor performance but disproportionally flock to funds with the best recent performance. This flow-performance relation incentives fund managers to compete in the short-term, and take myopic investment decisions at the expense of long-term performance (Lin and Kogan (2005)). I find that even though recent performance is important for investors, the effect of recent performance on fund flows is highly dependent on long-term performance. Specifically, long-term winners can afford a losing quarter, while long-term losers cannot increase their flows by boosting their recent performance. Hence, short-termism on behalf of fund managers may not be a competitively viable strategy, because fund investors see through attempts to boost short-term performance at the expense of long-term performance.

The findings of the first and second essay imply that consumers use a complex multi-attribute model when optimizing their fund choices. The value that investors attach to long-term performance and lottery-like returns may differ among different clientele. Accordingly, I show that preference for lottery-like payoffs only exist in retail funds, which predominantly cater to individual investors. Nevertheless, as long as long a sizable portion of investors consider these fund characteristics in their purchase decisions, we can observe the effect of both fund characteristics on fund flows.

APPENDIX

SAMPLE SELECTION

I start with the CRSP fund summary file, which includes yearly observations of fund characteristics except the returns and total net assets. The selection criterion is based on the fund objectives. There are two different fund objectives in this CRSP file after 1993: Strategic Insights (SI) objective codes, and Lipper objective codes. These codes cover two different time periods: the Strategic Insights objective code is available between 1993 and 1998; and the Lipper objective code exists after 1999. To determine the equity funds, we select the funds with the following codes.

(1) Strategic Insights codes (SI_OBJ_CD): AGG, GMC, GRI, GRO, ING, SCG

(2) Lipper objective codes (LIPPER_OBJ_CD): CA, I, G, GI, I, MC, MR, SG

Then I regroup the funds into 5 classes:

(1) Capital appreciation- LIPPER_OBJ_CD: CA

(2) Growth & Income- SI_OBJ_CD: GRI; LIPPER_OBJ_CD: GI

(3) Income- SI_OBJ_CD: ING; LIPPER_OBJ_CD: EI,I

(4) Growth- SI_OBJ_CD: AGG, GRO,GMC; LIPPER_OBJ_CD: MC, G

(5) Small Growth- SI_OBJ_CD: SCG; LIPPER_OBJ_CD: SG, MR

An objective code in its valid period may be missing for a fund in some years. If a code is missing in a particular year but is available for that fund in a later or earlier year, then we keep that observation and fill the code with its earlier or later classification. Funds with "*index fund flag*" equal to "D” are further deleted from the sample because they are pure index funds. However, this flag is available only after June 2008. Hence, I also make a name search to identify index funds. Following Gil-Bazo and Ruiz-Verdύ

(2009), I code a fund as an index fund if its name contains any of the following strings: "Index", "Idx", "Ix", "Indx", "NASDAQ", "Nasdaq", "Dow", "Mkt", "DJ", "S&P 500", and "BARRA". Evans (2010) points out a strategy in which fund families start multiple new funds and open some of them to the public at the end of an evaluation periods and terminate the others. This strategy creates an upward bias in fund returns. To address this incubation bias, we exclude the fund with missing CRSP names and that are less than 2 years old. CRSP calculates returns from one non-missing net asset value (NAV) to another. Hence, when I observe a missing NAV, I delete the next month return. Elton, Gruber, and Blake (2001) find an upward bias in the returns of small funds. Moreover, the incubation bias documented in Evans (2010) tends to be concentrated in small funds. Therefore, we exclude the funds with total net asset below $5 million in the previous month to partially address these concerns

REFERENCES CITED

Albuquerque, Rui, 2012, Skewness in stock returns: Reconciling the evidence on firm versus aggregate returns, *Review of Financial Studies*, Forthcoming.

Alexander, Gordon J, Gjergji Cici, and Scott Gibson, 2007, Does motivation matter when assessing trade performance? An analysis of mutual funds, *Review of Financial Studies* 20, 125-150.

Åstebro, Thomas, José Mata, and Luís Santos-Pinto, 2009, Preference for skew in lotteries: Evidence from laboratory, Working Paper, HEC Paris.

Bailey, Warren, Alok Kumar, and David Ng, 2011, Behavioral biases of mutual fund investors, *Journal of Financial Economics*, Forthcoming.

Baker, M., 1998, Fund managers' attitudes to risk and time horizons: the effect of performance benchmarking, *The European Journal of Finance* 4, 257-278.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427-446.

Barber, Brad M., Terrance Odean, and Lu Zheng, 2005, Out of sight, out of mind: the effects of expenses on mutual fund flows, *Journal of Business* 78, 2095-2120.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161-199.

Barberis, Nicholas, and Ming Huang, 2008, Stocks as lotteries: The implication of probability weighting for security prices, *American Economic Review* 98, 2066-2100.

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

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269-1295.

Bilias, Yannis, Dimitris Georgarakos, and Micheal Haliassos (2009), Portfolio inertia and stock market fluctuations, *Journal of Money, Credit and Banking* 42, 715-742.

Boyer, Brain H., and Keith Vorkink, 2011, Stock options as lotteries, Working Paper, Brigham Young University.

Boyer, Brain H., Todd Mitton, and Keith Vorkink, 2010, Expected idiosyncratic skewness, *Review of Financial Studies* 23, 169-202.

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.

Brunnermeier, Markus K., Christian Gollier, and Jonathan A. Parker, 2007, Optimal beliefs, asset prices and the preference for skewed returns, *American Economic Review* 97, 159-165.

Bushee, Brian J, 1998, The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review* 73, 305-333.

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

Cashman, George D., Daniel N. Deli, Federico Nardari, and Sriram V. Villupuram, 2006, Investors do respond to poor mutual fund performance: evidence from inflows and outflows, Working Paper, Arizona State University.

Chen, Joseph, Harrison Hong, Mind Muang, and Jeffrey D. Kubik, 2004, Does find size erode mutual fund performance? The role of liquidity and organization, *American Economic Review* 94, 1276-1302.

Chen, Joseph, Samuel Hanson, Harrison Hong, and Jeremy C. Stein, 2008, Do hedge funds profit from mutual-fund distress, Working paper, Harvard University.

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

Chevalier, Judith A., and Glenn D. Ellison, 1998, Career concerns of mutual fund managers, *The Quarterly Journal of Economics* 113, 389-432.

Clifford, Christopher P., and Bradford J. Jordan, 2011, The average investor does not own the average funds: Implications for mutual fund research, Working Paper, Universtiy of Kentucky.

Conrad, Jennifer S., Robert F. Dittmar, and Eric Ghysels, 2009, Ex-ante skewness and expected stock returns, Working paper, University of Michigan.

Cooper, Michael J., Huseyin Gulen, and P. Raghavendra Rau, 2005, Changing names with style: Mutual fund name changes and their effects on fund flows, *Journal of Finance* 60, 2825-2858.

Coval, Joshua D., and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479-512.

Del Guercio, Diane, and Paula A. Tkac, 2008, Star power: The effect of Morningstar ratings on mutual fund flows, *Journal of Financial and Quantitative Analysis* 4, 907-936.

Duffee, Gregory R., 1995, Stock returns and volatility, *Journal of Financial Economics* 37, 399–420.

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, 2001, A first look at the accuracy of CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund database, *Journal of Finance* 56, 2415–2430.

Elton, Edwin J., Martin J. Gruber, and Jeffrey A. Busse, 2004, Are investors rational? Choices among index funds, *Journal of Finance* 59, 261-287.

Evans, Richard B, 2010, Mutual fund incubation, *Journal of Finance* 65, 1581-1611.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fishbein, Martin, and Icek Azjen, 1975, *Belief, Attitude, Intention, and Behavior* (Addison-Wesley Publishing, Reading, MA).

Frazinni, Andrea, and Owen A. Lamont, 2008, Dumb money: Mutual fund flows and the cross-section of stock returns, *Journal of Financial Economics* 88, 299-322.

French, Kenneth R., 2008, Presidential address: The cost of active investing, *Journal of Finance* 58, 1537-1573.

Gallaher, Steven T., Ron Kaniel, and Laura Starks, 2008, Advertising and mutual funds: From families to individual funds, Working Paper, University of Texas.

Gaspar, Jose´-Miguel, Massimo Massa, and Pedro Matos, 2005, Shareholder investment horizons and the market for corporate control, *Journal of Financial Economics* 76, 135-165.

Gaspar, José-Miguel, Massimo Massa, and Pedro Matos, 2006, Favortisim in mutual fund families? Evidence on strategic crsoo subsidization, *Journal of Finance* 61, 71-103.

Gil-Bazo, Javier, and Pablo Ruiz-Verdύ, 2009, The relation between price and performance in the mutual fund industry, *Journal of Finance* 64, 2153-2182.

Goetzmann, William N., and Nadav Peles, 1997, Cognitive dissonance and mutual fund investors, *Journal of Financial Research* 2, 145-158.

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

Harvey, Campbell R., and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263-1295.

Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser, 1992, Nonrational actors and financial market behavior. *Theory and Science* 31, 257-287.

Hong, Harrison, and Jeremy C. Stein, 2003, Differences of opinion, short-sales constraints and market crashes. *Review of Financial Studies* 16, 487-525.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang, 2011, Risk shifting and mutual fund performance, *Review of Financial Studies* 24, 2575-2616.

Huang, Jennifer, Kelsey D. Wei, and Hong Yan, 2007, Participation costs and the sensitivity of fund flows to past performance, *Journal of Finance* 52, 1273-1311.

Huang, Jennifer, Kelsey D. Wei, and Hong Yan, 2012, Investor learning and mutual fund flows, Working Paper, University of Texas.

Investment Company Institute, 2011, Investment Company Fact book: A review of trends and activity in the investment company industry, Washington DC.

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.

Ivanov, V. and Kilian, L. 2005, A Practitioner's guide to lag-order selection for vector autoregressions. London Centre for Economic Policy Research.

Jain, Prem C., and Joanna Shuang Wu, 2000, Truth in mutual fund advertising: Evidence on future performance and fund flows, *Journal of Finance* 55, 937-957.

Jin, Li, and Leonid Kogan, 2007, How does investor short-termism affect mutual fund manager short-termism, Working Paper, Harvard Business School.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983-2011.

Keswani, Aneelm and David Stolin, 2008, Which money is smart? Mutual fund buys and sells of individual and institutional investors, *Journal of Finance* 63, 85-118.

Khan, Mozaffar, Leonig Kogan, and Geroge Serafim, 2011, Mututal fund trading pressure: Firm-level stock price impact and timing of SEOs, *Journal of Finance*, Forthcoming.

Kumar, Alok, 2009, Who gambles in the stock market? *Journal of Finance* 64, 1889-1933.

Lanchaster, Kelvin J., 1966, A new approach to consumer theory, *Journal of Political Economy* 74, 132-157.

Lynch, Anthony W., and David K. Musto, (2003), How investors interpret past returns?, *Journal of Finance* 58, 2033-2058.

Meulbroek, Lisa K., Mark L. Mitchell, J. Harold Mulherin, Jeffry M. Netter, and Annette B. Poulsen, 1990, Shark repellents and managerial myopia: an empirical test, *Journal of Political Economy* 98, 1108-1117.

Mitton, Todd, and Keith Vorkink, 2007, Equilibrium underdiversification and the preference for skewness, *Review of Financial Studies* 20, 1255-1288.

Mullainathan, Sendhil, 2001, Thinking through categories, Working Paper, MIT.

Nanda, Vikram K., Z. Jay Wang, and Lu Zheng, 2009, The ABCs of mutual funds: On the introduction multiple share classes, *Journal of Financial Intermediation* 18, 329-361.

Nanda, Vikram K., Z. Jay Wang, and Lu Zheng, 2004, Family values and the star phenomenon strategies of mutual fund families, *Review of Financial Studies* 17, 667-698.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

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

Pastor, Lubos, and Robert F. Stambaugh, 2002, Mutual fund performance and seemingly unrelated assets, *Journal of Financial Economics* 61, 315-349.

Peterson, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, 453-480.

Pollet, Joshua M., and Mungo Wilson, 2008, How does size affect mutual fund behavior?, *Journal of Finance* 63, 2941–2969.

Pomorski, Lukasz, 2004, Style investing: Evidence from mutual fund flows, Working Paper, University of Toronto.

Pontiff, Jeffrey, 1996, Costly arbitrage: Evidence from closed-end funds, *Quarterly Journal of Economics* 111, 1135-1151.

Sapp, Travis, and Ashish Tiwari, 2004, Does stock return momentum explain the "smart money" effect?, *Journal of Finance* 59, 2605-2622.

Sawicki, Julia, and Frank Finn, 2002, Smart money and small funds, *Journal of Business Finance and Accounting* 29, 825-846.

Scharfstein, David S., and Jeremy Stein C., 1990. Herd behavior and investment. *American Economic Review* 80 (3), 465–479.

Sensoy, Berk A., 2009, Performance evaluation and self-designated benchmark indexes in the mutual fund industry, *Journal of Financial Economics* 92, 25-39.

Sialm, Clemens, and T. Mandy Tham, 2011, Spillover effects in mutual fund companies, Working Paper, University of Texas.

Singletion, J. Clay, and John Wingender, 1986, Skewness persistence in common stock returns, *Journal of Financial and Quantitative Analysis* 21, 335-341.

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

Spiegel, Matthew, and Hong Zhang, 2010, Mutual fund flow timing and external growth, Working Paper, Yale School of Management and Insead.

Teo, Melvyn, and Sung-Jun Woo, 2001, Persistence in style-adjusted fund returns, Working Paper, Harvard University.

Tversky, Amos, and Daniel Kahneman, 1992, Advance in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* 5, 297-323.

Wahal, Sunil, and Albert (Yan) Wang, 2011, Competition among mutual funds, *Journal of Financial Economics* 99, 40-59.

Wermers, Russ, 2004, Is money really "smart"? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence, Working Paper, University of Maryland.

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