



**UNIVERSITY OF GOTHENBURG**  
**SCHOOL OF BUSINESS, ECONOMICS AND LAW**

Master Degree Project in Finance

# **Competition in the American Mutual Fund Industry**

An empirical study

Farhad Miri

Supervisor: Martin Holmén  
Master Degree Project No. 2014: 117  
Graduate School

## **Abstract**

This thesis investigates and compares the relationship between the inflow of new investment into open-end equity U.S. mutual funds and their historical performance among the top performer funds. Using a piecewise linear regression and applying the Fama and MacBeth (1973) two stages estimation method on the fund data over the period between January 2004 and December 2014, it was found that the level of convexity within top performer is more extreme than what is usually observed as a convexity in the flow-performance relation among the whole industry players. The difference is irrespective of the performance measurement and is both statistically and economically significant. The results obtained suggest that the competition among the mutual funds is not just about being better than average but is rather about winning the “competition”. Fund managers can achieve marked additional inflow related to their peers by securing their position among top 10% of the industry in terms of performance. A positive significant relation between Morningstar rating and fund flow, was also documented.

**JEL classification:** G23, C23

**Key Words:** Mutual Funds; Flow-Performance relationship; Convexity; Competition; Morning Star; Panel Data; Piecewise linear regression

### **Acknowledgements**

I would first like to thank my supervisor, Professor Martin Holmen for patient guidance, inspiration and recommendations he has provided throughout this thesis. I want to thank Senior Lecturer Evert Carlsson who provides me the access to the Morningstar database.

## Contents

1. Introduction.....	5
2. Literature Review.....	9
3. Data.....	11
3.1 Data Selection .....	11
3.2 Data cleaning.....	11
3.3 Outliers and Survivorship bias .....	12
3.4 Variable Definitions.....	12
3.4.1 Flows .....	15
3.4.2 Performance.....	16
3.4.3 Fees .....	17
3.4.4 Riskiness.....	17
3.4.5 Objective Mean Flow .....	17
3.5 Visual Analyze of Return, Risk, Fee and Rating.....	19
4. Methodology and Econometric Specification.....	24
5. Results and Discussion .....	26
5.1 Baseline Specification .....	26
5.2 Alternative Specification.....	30
5.3 Origin of convexity .....	33
6. Summary and Conclusion .....	36
References: .....	39

# 1. Introduction

The primary objective of this thesis is to explore the relationship between inflow of new investment into open end equity U.S mutual fund and their historical performances. There are numerous papers that address issue of flow-performance relationship (see e.g. Del Guercio & Tkac (2002) and Ippolito (1992)). Many of these studies attempt to estimate the general shape of this relationship by addressing the difference between the best and worst performers (see e.g. Sirri and Tufano (1998) and Chevalier and Ellison (1997)). Nevertheless, few studies investigate how the relationship differs among the best performers. I contribute with new results, by first estimating the general shape of the relationship using the latest dataset available for U.S equity mutual fund industry (time period between January 2004 and December 2013), and secondly explore the relationship for the top performers to investigate whether there exists a marked difference among the top performers. In particular I am interested to interpret the “competition” among mutual funds to attract new inflow. This thesis investigates the impact of fund’s position in the market on the subsequent flow to answer the question whether the competition for mutual funds means they have to just perform better than average or they have to compete against their peers to win the competition.

Investment institutions in general and mutual funds in particular have been developed noticeably over the past decades. The development in the mutual fund industry has lead financial scholars to study the parameters considered by investors in their fund selection process. Most of the literature studying mutual fund industry concur that investors prefer the funds with better performance. One of the possible explanations for this behavior could be the fact that investors hope mutual funds exhibit persistence in performance over time. Brown and Goetzmann, (1995) and Ibbotson and Goetzmann (1994) provide evidence of performance persistence among high performers. Most of the literature (for instance Ippolito (1992); Sirri and Tufano (1998)) coincide in stating that investors respond to fund’s performance disproportionately i.e. they do not sell the past loser at the rate that they buy past winners. According to the literature there are two main explanations for this convexity in flow-performance relationship. On the one hand Lynch and Musto (2003) interpret this phenomenon through the role of investment advisors to replace

fund's management or investment policy. On the other hand in the last few years a series of studies has focused on the role of searching cost and mutual fund's fee on the convexity of flow performance relationship. Navone et al. (2012) find evidence that marketing fees have a positive impact on flow-performance relationship convexity. Moreover, Berk and Green (2004) and Huang et al. (2011) provide empirical evidence that mutual fund's non-marketing fees are negatively related to the sensitivity of flow to the fund's historical performance.

The importance of flow-performance relationship stems from mutual fund managers receiving management fees as a percentage of assets under their management. Chevalier and Ellison (1997) show that this compensation structure induces an implicit incentive for fund's managers to change the riskiness of their portfolio. This could turn into an agency conflict between investors and fund managers. Therefore understanding the flow performance relationship in more detail should persuade regulators to propose modifications in law for more transparency in terms of financial information provided by mutual funds to investors.

There is contradictory evidence regarding the persistence of mutual fund's return particularly among funds with high performance. Hendricks et al. (1993) provide evidence that the worst performing funds repeat their performances. Although, Malkiel (1995) find no evidence for consistency in fund returns, other studies, Brown et al. (1992) and Grinblatt and Titman (1992) find persistence among winners i.e. high performers persistence.

The main question and primary testable hypothesis this thesis attempts to answer is whether the convexity of flow performance relationship differs noticeably among the top mutual fund's performers in the open end U.S equity mutual fund industry. To this end, first the flow performance relationship is estimated and a comparison between the top and bottom performance quintile is done. Second, a more detail investigation is done for those funds that are among top performers.

Chevalier and Ellison (1997) use a "Semi-parametric" model and find evidence that fund's age affect flow-performance relationship. I also test this hypothesis by comparing the flow performance relationship among different age categories. Furthermore, Keswani and Stolin (2012) show that different type of investors react differently to fund's performance. 22% of the funds in the dataset are the funds designed for institutional investors and the rest are aimed at

non-institutional investors. Flow-performance relationship is analyzed and compared separately for these two types of fund to investigate whether they have noticeable different flow-performance relationship.

My sample consists of 1292 U.S equity mutual funds 2004 to 2013. I first document that the relation between new inflow into the fund and historical performance of the fund is a convex relation. Moreover I find that the convexity of the flow-performance relationship increases by the level of fund's performance. The better performance the fund has relative to its peers the more it would benefit from flow performance relationship. Besides, I find younger and smaller funds benefit more from the flow-performance relationship and the results also suggest that investors are sensitive to the fund's riskiness measured as standard deviation of fund's return. Finally, searching for the origin of convexity in the relation, the results suggest that retail investors are more extreme in terms of investing in past winners while failing to flee from past losers at the same rate.

Earlier studies attempts to predict the shape of flow performance relationship in general. I extend these studies by examining and comparing the flow-performance relationship among top performers. I find that the level of convexity within top performer is more extreme than what is usually observed as a convexity in the flow-performance relation among the whole industry. The performances of the funds within top 10% are rewarded more strongly compared to the funds within top 20%. The difference is significant both statistically and economically. The results point out that the tournament in the mutual fund industry is about winning the competition by being among highest possible position, rather than performing better than average.

I believe understanding the way that investors response to fund' performance could be served as guide line for regulators and fund managers as well as investors themselves. Since the flow-performance relationship determines by the level of investor's sophistication and sensitivity, regulators can use this knowledge to improve investor's performance. Moreover, these results suggest to the fund managers that only performing better than average does not result in additional inflow and if they wish to attract marked additional inflow related to their peers they need to secure their position among top 10% of the industry in terms of performance.

The rest of the thesis is structured as follows. The next section reviews the previous literature regarding the flow-performance relationship. The following section describes the data selection and data cleaning process and defines the variables to be analyzed. It also outlines the methodology. I then continue by reporting and discussing the main results. Finally, the last section concludes and discusses the implications of the results.



## 2. Literature Review

Numerous studies have analyzed the relationship between fund's historical performance and the flow of new investment into the mutual fund (see e.g. Ippolito (1992); Chevalier and Ellison (1997); Ferreira et al (2012); Del Guercio and Tkac (2002)). They attempt to estimate the shape of the flow-performance relationship by using different models and different sets of control variables. Most studies conclude that there exists a strong positive relationship between fund's past performance and flow (e.g. Sirri and Tufano (1998)). Most of these literatures note that the relation is nonlinear so they try to allow for this nonlinearity. Chevalier and Ellison (1997) use a "Semi-parametric" model to estimate the shape of this relation. They find a convex relationship between flow and performance and argued that considering structure of management fees in this industry, the relation generate an implicit incentives for fund managers to increase or decrease the riskiness of their portfolio. They compare the funds' holding between September and December and show that the fund managers alter their portfolios riskiness according to this incentive.

Using a piecewise linear regression Sirri and Tufano (1998) confirm that the relation between performance and inflow is highly dependent on level of fund's performance. They find that investors response asymmetrically to mutual fund's performance. While there is a slightly negative relation between past performance and new inflow for the worst mutual fund (based on historical performance), there is a substantial positive relationship between past performance and new inflow for the best mutual fund performers.

Although, most of the papers focus on mutual fund industry in US market, there are some papers which investigate the relation in other markets. Alves and Mendes (2011) study the investor's reactions to fund's past performance using a sample of Portuguese open-end equity funds and find no evidence of convex relationship. Ferreira et al. (2012) did a cross country study on mutual funds flow in 28 countries and compare the flow-performance relationship in different countries. They find evidence that the convex relationship found in US mutual funds industry do not apply universally and the convexity level is higher in more developed countries. They

explain this cross-country differences by arguing that investors in more developed countries are more sophisticated and facing lower participation costs.

Most of the literatures confirm that the flow performance relationship differs for top performers from bottom performers. However, to my knowledge there are few studies that investigates to what extent the flow performance relationship differs among top players. Kempf and Ruenzi (2008) analyze the flow performance relationship focusing on the fund's position within its family. They find evidence that additional to the fund's position in the market fund's relative position in its family also matters. More importantly they show that if a fund improves its position and reaches to the top 10% in its family it makes a huge difference for the fund in terms of subsequent inflow. However, their result regarding the implication of competition for the top players limits to the fund's position in its family. The novelty of this thesis is that it studies the flow performance relationship for top players in more detail. It investigates the impact of fund's position in the market on the subsequent flow to discover whether the competition in mutual fund industry means that a fund should just perform better than the average or it is about winning the competition.

## 3. Data

### 3.1 Data Selection

To investigate the relation between Cash-flow and Performance of mutual funds, data on mutual fund sizes and returns are drawn from the Morningstar Inc. Morningstar is an investment research platform, which provides data on approximately 433,000 investment offerings including stocks, mutual funds and other similar vehicles. I obtain data on open end U.S equity mutual fund net asset values (NAV), total net asset (TNA), fund's relevant fees and other classification variables such as fund age and fund prospectus objective category.

### 3.2 Data cleaning

Information on actively managed U.S equity open end funds has been drawn from the database for the time interval between January 2004 and December 2013. Most of the funds have more than one share class, which Morningstar shows as different funds. However, these different share classes have the same managers and same holdings and just differ in their management fees. These funds have the same return before management expenses and other fees and if one treats them separately there would be multi counting problems. Therefore, following Ferreira et al. (2012) I begin by eliminating the multiple share classes and just the funds with primary share classes<sup>1</sup> (oldest share class) have been kept in the sample. Some of the funds have been closed to investors. In this case since the performance of the funds could not affect the fund's cash flow I consider them to be irrelevant for the flow performance relationship analysis. Moreover, for the same reason those funds, which are closed to new investor, have been removed from the sample. Chevalier & Ellison (1997) argue that these funds and also the funds that are known as index funds and the funds of the funds might have different flow nature. Therefore they have been removed from the sample. Finally, few funds in the dataset have been merged. Since these mergers results in a misleading cash-flow these funds have also been removed from the dataset. McDonald (1974) notes that the funds with Growth objective in their classifications have a desire to change the riskiness of their portfolio regarding their performance. This property makes the

---

<sup>1</sup> The primary share class is usually defined as the class with the highest total net assets (TNA).

funds with Growth objective a suitable choice when studying the flow-performance relationship. Sirri and Tufano (1998) also consider the funds with aggressive growth and growth and income in their study. In the same manner I restrict the sample to the mutual funds within these three prospectus objective category; Growth, Aggressive Growth and Growth and Income.

After implementing all these criteria 1292 funds have been selected for the time period between January 2004 and December 2013.

### **3.3 Outliers and Survivorship bias**

There were some data points, which were suspicious regarding the annual flow and return. All the questionable data points have been checked and compared with Bloomberg and any doubtful data points were reported to Morningstar and/or corrected with the Bloomberg dataset.

One of the main concerns regarding the study of mutual funds industry is sample selection bias particularly when one attempt to collect information about the fund's performance and flow. If one excludes the defunct funds from the sample, then there would be survivorship bias problem (Brown et al. (1997)). In my dataset most of the mutual funds were active in the time period of study. Morningstar also provide information for the funds that went out of existence. Following Goetzmann and Peles (1997) for each dead fund the flow for the year after its death is set to -100%. Robustness tests show that omitting portfolios that have been dissolved during this period does not affect the results.

### **3.4 Variable Definitions**

To analyze the flow-performance relationship I took the relevant financial information as well as classification variables for the open end U.S equity mutual funds for the period between January 2004 and December 2013. The Financial data points consist of:

*Total Net Asset (TNA)* which is the total dollar value of the fund portfolio; TNAs have been drawn on a monthly basis and is used to form the annual growth rate of new cash flow into the fund's portfolio.

*Net Asset Values (NAV)*; this is per share dollar value of the fund's portfolio which is used to calculate the return of funds as a measure of fund's performance. Furthermore, all the fund's distributions such as capital gain and dividends have been captured wherever applicable.

Moreover I collect the management fees including *Annual Net Expense Ratio* and commission fee (*Front Load Fees*) for each fund in the sample. The expense ratio is the actual fee that is charged by the fund each fiscal year. It includes all the asset-based costs brought by the fund and shows the percentage of assets deducted yearly for the fund expenses. Front load fee is the initial sale charge that is deducted from each investment in the funds and is expressed as a percentage of investment.

Finally, other funds classification variables such as *Age*, *Prospectus Objective Category* and *Rating* are drawn from the Morningstar. Prospectus objective specifies fund's investment goals, based on the expressing in a fund's prospectus. My dataset consists of fund with these three classification; *Aggressive Growth*, *Growth* and *Growth and Income*.

*Aggressive growth*; includes funds that look for rapid growth of capital and mainly invest in small or emerging market growth companies.

*Growth*; includes funds that their primary concern is growth of capital. Therefore they primarily invest in equity securities.

*Growth and income*; includes funds that their objectives are equally growth of capital and current income. (Morning Star Glossary, 2014)

Since 1985, Morningstar introduced a fund ranking measurement called "Star Rating". This is a normally distributed yearly rating using the scale of one to five stars for each fund based on fund historical performance. (Morning Star Glossary, 2014) Del Guercio and Tkac (2002), explicitly test whether Morningstar rating has any influence on investment allocation decision of mutual fund investors. Through an event study over 100,000 Morningstar rating change, they find substantial abnormal flow following rating upgrades and vice versa. I control for the impact of Morningstar rating on the flow performance relationship by including the "Star Rating" in the model.

According to Morningstar definition a fund is defined as an Institutional fund if it has one of the following qualifications;

- a. “It requires \$100,000 or more as a minimum initial purchase
- b. It has the word "institutional" in its name
- c. It states in its prospectus that it is designed for institutional investors or those purchasing on a fiduciary basis”. (Morning Star Glossary, 2014)

Around 78% of the funds in the data set are non-institutional funds. Keswani and Stolin (2012), note that there exists a marked difference in reaction to fund performance between different type of institutional and retail investors. Therefore, this categorization is used as a control variable to compare the flow performance relationship between different types of funds. Based on the fund’s inception date the fund age has been calculated. Due to the fact that very young funds could have a noisy flow, those with less than 1 year age have been removed from the dataset (Chevalier & Ellison, 1997).

Previous literature notices that mutual fund’s size affect flow-performance relationship. Del Guercio and Tkac (2008) suggest that this disclose the significance of agency relationship and client servicing. However, Jain and Wu (2000), Sapp and Tiwari (2004) and Alves and Mendes (2011) argue that the size of fund could be a reflection of fund’s reputation and visibility (Ballester, 2014). I include the natural logarithm of fund’s total asset as a measure of fund’s size in my model to control the impact of fund’s size on flow-performance relationship.

Table 1 presents a descriptive statistics for the primary explanatory variables used in this study.

**Table 1: Descriptive Statistics for primary explanatory variables**

	Mean	SD	Min	Max	Count
Net Assets Value (\$)	19.26	28.04	1.64	807.83	9434
Capital Gain (\$)	0.46	1.14	0.00	31.93	9434
Dividend (\$)	0.09	0.25	0.00	10.57	9434
Total Net assets (TNA) (million \$)	1160	4470	0.01	91400	9434
Log of TNA	18.92	2.13	9.21	25.24	9434
Annual Net Expense Ratio (%)	1.16	0.42	0.00	4.94	9434
Front Load Fee (%)	1.82	2.59	0.00	5.75	9434
Total Fee (%)	1.42	0.62	0.00	5.30	9434
Age (Year)	14.99	13.65	1.00	89.42	9434
Institutional fund	0.22	0.41	0.00	1.00	9434
Rating	3.04	1.02	1.00	5.00	8646
Observations	9434				

The sample includes 1292 open end equity mutual funds through the January 2004 and December 2013. There are 9434 fund-year observations in the sample. Following Sirri and Tufano (1997) "Total Fee" estimated as the sum of "Annual Net Expenses Ratio" and the amortized "Front Load Fee". The front load fee is amortized over the seven years period. Seven years is the average holding period of mutual funds (Investment Company Institute, FactBook (2014)). For each variable the mean, standard deviation and max and min are reported.

### 3.4.1 Flows

Flow is defined as the net growth rate of new inflow into the fund. Here the assumption is that all the dividend and capital gains are reinvested immediately, therefore the flows reflect only the growth rate due to new money invested in the fund and not due to the dividends and capital gains earned on the assets under management. Following Sirri and Tufano (1998) and Chevalier and Ellison (1996) the flow of new money into the fund is defined as the growth rate of total net asset under the management of the fund between the year  $t-1$  and year  $t$  which has not occurred due to the return of the fund in the year  $t$ . i.e its net of internal growth. Formally the flow is:

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

Where,  $TNA_{i,t}$  is the total net asset value of fund  $i$  at the year  $t$  and  $R_{i,t}$  is funds  $i$ 's raw return over the year  $t$ , both reported in U.S dollar. This formulation assumes that flows occur yearly. However, Sirri & Tufano (1998) show that the flow-performance relation is irrelevant to the assumption about the schedule of flow occurrence. Funds are divided into 100 equally group based on the annual flow. To prevent the results from being affected by extreme values and following the Keswani & Stolin (2012) approach, I drop the top and bottom groups, i.e. the top and bottom 1%.

### 3.4.2 Performance

Previous literature use different measures of fund performance to investigate the flow-performance relationship (For instance fund's Historical Raw Returns, Return net of expenses, Jensen's Alpha and Four Factor Alpha). However, Sirri and Tufano (1998) show that irrespective of performance measurement the flow-performance convexity exists i.e. the flow-performance relationship is independent to any specific measure of performance. Moreover, they notice that among the different types of the performance measures, the investors most commonly have access to "Historical Raw Return". Therefore I assume that the investors might respond to these measures of performance more appropriately and use the Raw Return as the base measure of performance throughout this thesis and Return Net of Expenses as an alternative measurement. The primary raw return used in the baseline specification is one year raw return; however I use two, three and five year raw return in alternative specifications.



### 3.4.3 Fees

Mutual funds charge investors two general types of fees. One is an upfront fee which is an initial sales charge paid by investors when the purchase is done. The other is ongoing fees like the annual net expense ratio that is paid annually as a percentage of assets under fund management. The expense ratio includes management fees, administrative fees, operating costs, 12b-1 fees and all other asset-based costs paid by the fund except the initial or deferred sales charges and portfolio transaction fees. 12b-1 fee is a fee to compensate mutual fund's distribution costs and is usually paid as a commission to brokers (Morning Star Glossary, 2014).

Following Sirri and Tufano (1998) I calculate the total fee by summing up these fees. Considering the average holding period of the equity fund which is seven years, the upfront load is amortized over this period and is added up to the annual net expense ratio (Investment Company Institute, FactBook (2014)). Formally the total fee is:

$$Total\ fee = upfront\ fee \times \left(\frac{1}{7}\right) + Annual\ net\ expense\ ratio$$

This approach results in a rough estimation of the total fee but since I aim to investigate the impact of fees on the flow performance relationship it gives better insight than the component of the fees separately.

### 3.4.4 Riskiness

To investigate the impact of fund's portfolio riskiness on flow-performance relationship, standard deviation of fund's monthly return over last year return (previous 12 months) is calculated and used as an explanatory variable in the main specification. In an alternative specification instead of one year return, two, three and five year returns are used and the standard deviation is calculated accordingly as the fund's monthly return over two, three and five year period, respectively.

### 3.4.5 Objective Mean Flow

Since part of the inflow into the fund might come from increasing the flow into the whole market we control for this by including the average of flow into each objective category and for each

year. Table 2 shows the descriptive statistics of flow, performance, riskiness and object mean flow.

**Table 2: Descriptive statistics of flows, performance, riskiness and objective mean flow**

	Mean	SD	Min	Max	Count
Annual Flow	0.1502	0.8014	-0.6419	9.1789	9434
Raw return	0.0999	0.2177	-0.6504	1.6561	9434
Riskiness	0.0449	0.0190	0.0033	0.1911	9434
Objective Mean Flow (Aggressive Growth)	0.1627	0.0891	0.0535	0.3392	298
Objective Mean Flow (Growth)	0.0501	0.1141	-0.0952	0.2753	7560
Objective Mean Flow (Growth and Income)	0.1090	0.0677	0.0145	0.2506	1576
Observations	9434				

The table shows the descriptive statistics of Annual flow, return, riskiness and objective mean flow. Annual flow is defined as  $(TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t}))/TNA_{i,t-1}$  where  $TNA$  is fund's total net assets and  $R_{i,t}$  is fund's one year raw return. Raw return is fund's one year raw return. Riskiness is standard deviation of fund's return over past twelve month return. Finally objective mean flow is defined as the average of flow into each objective category and for each year. For each variable the mean, standard deviation and max and min are reported.

Finally, to check for the existence of multicollinearity problems among fund's control variables, I generate a correlation matrix for independent variables. Table 3 presents the pairwise correlations among the control variables and the results indicate that most of the correlation coefficients are below 0.3. However, there is a high correlation between Log TNA and log Age (0.5) because older funds are usually bigger. Moreover, there is noticeably negative correlation (-0.39) between average objective category flows (CategoryFlow) and standard deviation of monthly return (Volatility). This could be due to the fact that as the volatility increases the investors are less willing to invest into mutual funds and as a result the average inflow into the whole market decreases.

Table 3: Pairwise correlations matrix among control variables

	1	2	3	4	5	6	
Log Age	1						
Log TNA	2	0.50***	1				
Total Fee	3	-0.05***	-0.29***	1			
Rating	4	-0.04***	0.16***	-0.29***	1		
Volatility (SD)	5	0.05***	-0.08***	0.04***	-0.10***	1	
CategoryFlow	6	-0.10***	0.02***	0.02**	0.07***	-0.39***	1

Table 3 presents the pairwise correlation matrix among fund's control variables. Control variables include: natural logarithm of fund Age in year (Log Age), natural logarithm of fund TNA as a measure of fund size (Log TNA), Total fee calculated as sum of annual net expense ratio and amortized front load fee, Morning star rating from the scale one to five (Rating), Standard deviation of monthly return over past 12 month return as a measure of fund's portfolio riskiness (Volatility) and average of flow into the each objective category of fund and for each year (CategoryFlow).

\*\*\* 1% Significance level

\*\* 5% Significance level

\* 10% Significance level

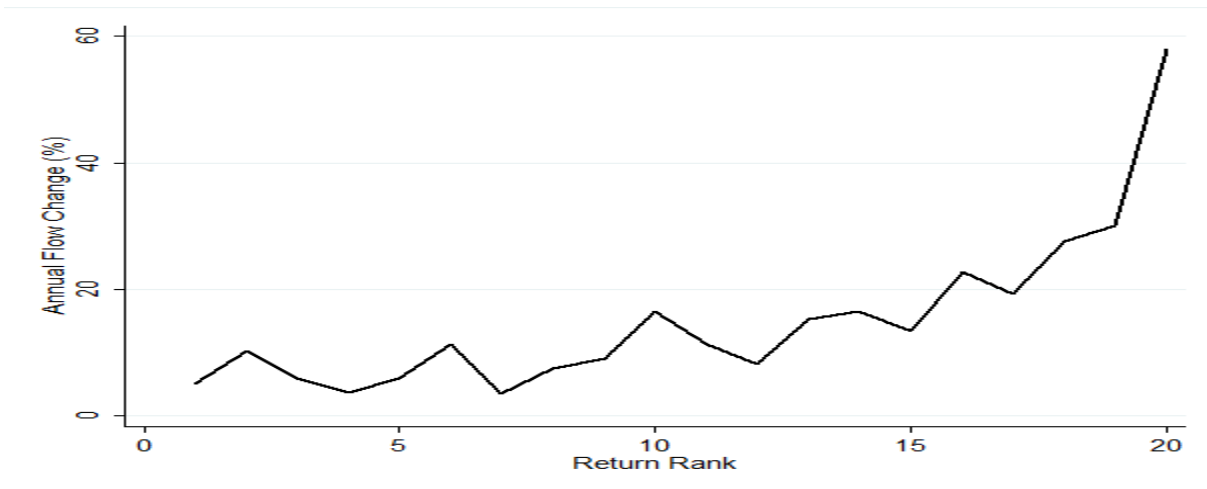
### 3.5 Visual Analyze of Return, Risk, Fee and Rating

To analyze better the relationship between flow and performance as well as other main explanatory variables (fund expenses, riskiness and rating), Figures 1 to 4 are plotted. For each year between 2004 and 2013 and for each prospectus objective category (Aggressive Growth, Growth and Growth and Income) the funds are divided and ranked into 20 equal groups. Based on their one year raw return (figure 1), standard deviation of monthly return over last year return (figure 2) and total fees (figure 3), the funds are ranked from one to twenty. Here rank 1 means the funds with worst performance, lowest total fees and lowest standard deviation of monthly return and accordingly for rank 20. Then for each of these twenty groups the average annual flow -over 10 years and three objective categories- for the next year is calculated and plotted against the funds that compromise each of these twenty groups. The figures 1 – 4 plot average net flow against last year raw return, standard deviation of monthly return total fee and rating respectively.

Figure 1 shows the relationship between flow and historical one year return which is quite considerable. For the worst 25% performers there exists almost no relation meaning there might be no penalty at all for funds with poor performance. Meanwhile, for the middle 50 % performers there is a slightly positive relation between past performance ranking and flow. And finally for the top 25% performers there is noticeable positive relation between growth of flow

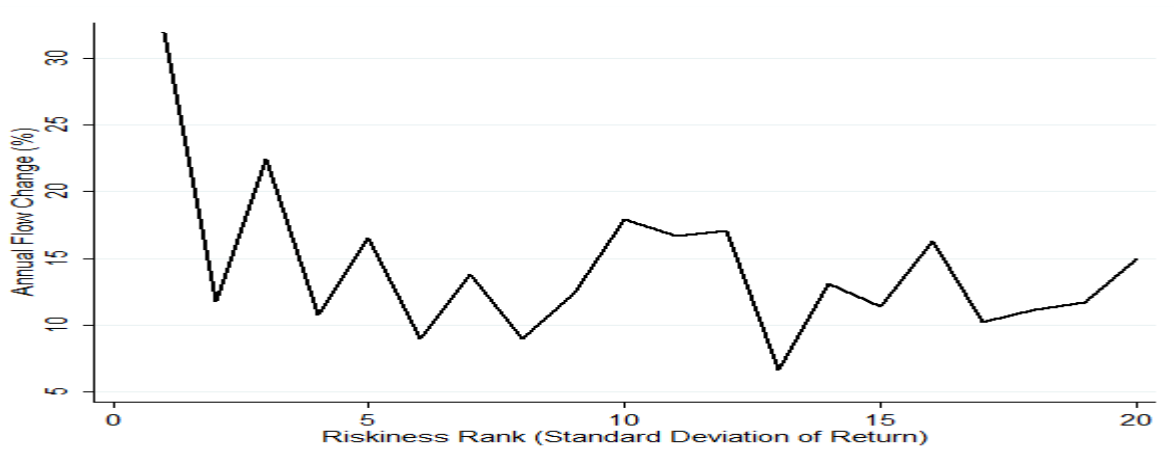
and last year raw return. The substantial positive slope for the top 5% performers indicates that improving performance ranking and reaching to the top position might awarded extensively. Some of earlier papers which study flow-performance relationship address a linear relation between growths of flow and performance of the funds (see e.g. Spitz (1970) and Smith (1978)). However, recent studies pronounce the non-linearity of the relationship which is confirmed by figure 1. (See e.g. Carhart (1994), Goetzmann and Peles (1997) and Sirri and Tufano (1998))

Figure 1: One year Raw Return and Flow



The figure shows the relationship between the fund historical performance ranking in a year and flow of the fund in the following year. For each year and each prospectus objective (Aggressive Growth, Growth and Growth and Income) funds are divided and ranked into 20 equal groups based on their one year raw return. Where rank 1 is the fund with worst performance in year t-1 and rank 20 is the fund with best performance in year t-1. Then the average annual growth rate of flow in year t calculated and plotted against the performance ranking.

Figure 2: Standard Deviation of one year Raw Return and Flow



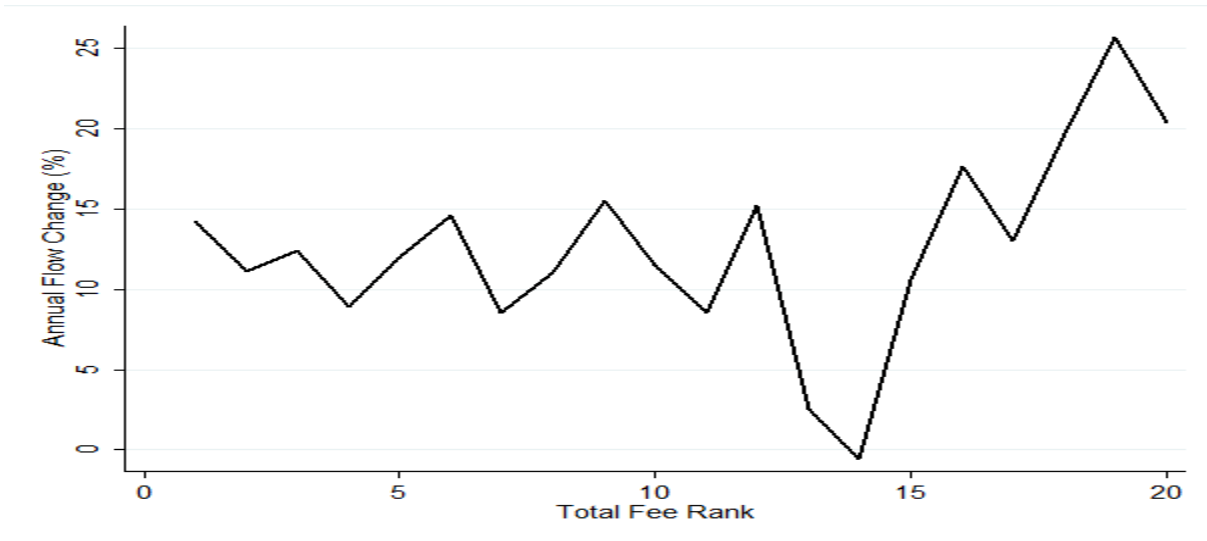
The figure shows the relationship between the fund standard deviation of return ranking in a year and flow of the fund in the following year. For each year and each prospectus objective (Aggressive Growth, Growth and Growth and Income) funds are divided and ranked into 20 equal groups based on their one year return SD. Where rank 1 is the fund with lowest SD in year t-1 and rank 20 is the fund with highest SD in year t-1. Then the average annual growth rate of flow in year t calculated and plotted against the SD ranking.

Figure 2 illustrates the relationship between flow and standard deviation of fund monthly return as a measure of fund riskiness. Standard deviation is calculated by using fund monthly return over last year return (past 12 months). Based on this measurement, funds are divided and ranked into 20 equal groups. For each group the average growth rate of fund annual flow is calculated and plotted against standard deviation ranking. A general negative relationship exists between annual flow and standard deviation of the fund’s portfolio for the bottom 40% which indicates the investors might penalize funds with higher riskiness. However, for the top 60% the relation is not clear.

Figure 3 shows the relation between total fee ranking and flow. The total fee is compromised by adding upfront load fees (purchasing commissions) and annual net expense ratio both as a percentage of assets under management. Considering the amortization of upfront fees, total fee is annual fee plus one seventh of upfront fees (Sirri & Tufano, 1998). The result suggests investors are sensitive to fund related fees. The figure suggests the funds with the highest fee ranking (the funds with highest percentage of total fee) attract highest annual flow. Sirri and Tufano (1998)

explain this through marketing costs and argue that the funds with higher fees have higher marketing effort; therefore they could benefit more from flow performance relation. The relation between fee rank and annual flow for the rest of the funds is unclear.

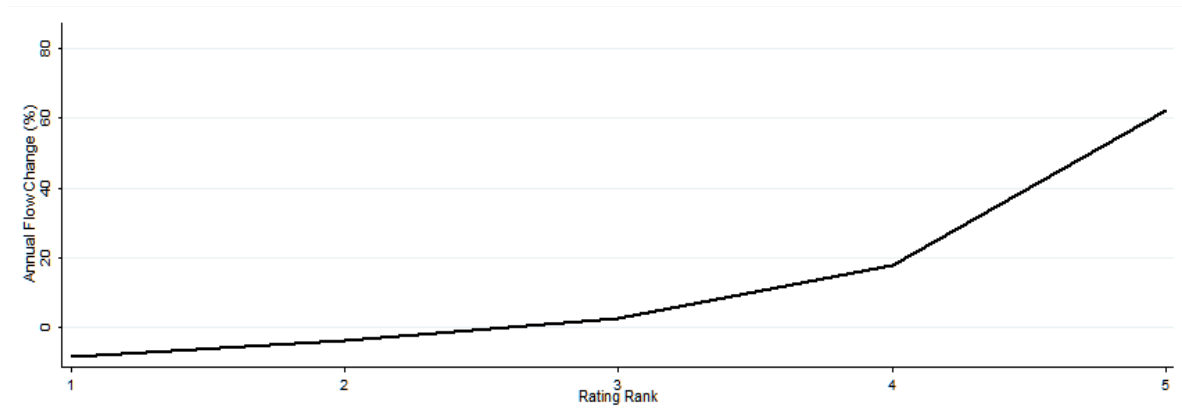
Figure 3: Total Fee and Flow



The figure shows the relationship between the fund total fee ranking in a year and flow of the fund in the following year. For each year between 2004 and 2013 and for each prospectus objective (Aggressive Growth, Growth and Growth and Income), funds are divided and ranked into 20 equal groups based on their total fee. Where rank 1 is the fund with lowest fee in year t-1 and rank 20 is the fund with highest fee in year t-1. Then the average annual growth rate of flow in year t calculated and plotted against the total fee ranking.

Finally, for each level of Morningstar rating (1 star to 5 stars) the average inflow is calculated for each year between 2004 and 2013 and for each prospectus objective category (Aggressive Growth, Growth and Income and Income). Figure 4 illustrates the relation between growth of flow and rating. Not surprisingly, the results show that funds with higher Morningstar rating are awarded with higher inflow rate.

Figure 4: Morningstar Rating and Flow



The Figure shows the relationship between the fund Morningstar Rating in a year and flow of the fund in the following year. For each year between 2004 and 2013 and for each prospectus objective (Aggressive Growth, Growth and Growth and Income the average annual growth rate of flow in year t calculated and plotted against the Fund rating in year t-1.

To sum up, figures 1 to 4 in general suggest that investors prefer higher return and rating along with lower volatility of return. However, since some of these explanatory variables might be related to each other, further investigation is needed in order to draw conclusions about the flow-performance relationship.

## 4. Methodology and Econometric Specification

I estimate the following model to estimate the relation between Flow and Performance:

$$\begin{aligned} Flow_{i,t} = & \beta_0 + \beta_1 LowPerformance_{i,t-1} + \beta_2 MiddlePerformance_{i,t-1} \\ & + \beta_3 HighPerformance_{i,t-1} + \beta_4 Volatility_{i,t-1} + \beta_5 Fees_{i,t-1} \\ & + \beta_6 CategoryFlow_t + \beta_7 Log TNA_{i,t-1} + \beta_8 Log Age_{i,t-1} + \beta_9 Rating_{i,t-1} \\ & + \varepsilon_{i,t} \end{aligned}$$

Where;

*Flow* represents the annual growth rate of fund's total net assets

*Volatility* measure by standard deviation of fund monthly return over last year return

*Fee* is the total fee calculated as the sum of annual expense ratio and one seventh of upfront loads.

*CategoryFlow* is the average of the inflow into the funds with same objective category

*Log TNA* measures the total net asset of the fund as a measure of fund size

*Log Age* is the natural logarithm of the fund age in year

*Rating* represents the Morningstar rating in the scale of one to five

And *Performance coefficients* are the fund return ranking which are calculated as follows

Following Sirri and Tufano (1998) I use a piecewise-linear regression, which allows the flow performance relationship to differ for different level of fund performance. To this end, I divide the funds based on their historical raw return into three main groups namely, *LowPerformance*, *MiddlePerformance* and *HighPerformance*. In the manner of Sirri and Tufano (1998) and Ferreira et al. (2012) performance groups are defined as follows:

For each year between 2004 and 2013 and for each prospectus objective category funds are ranked based on their annual raw return. Each fund has a rank between 0 to 1, where zero and one presents the poorest and best performers of the year in each objective category, respectively.

Then *Low*, *Middle* and *High performance* groups are defined as:



$$LowPerformance_t = \min(Rank_t, 0.2)$$

$$MiddlePerformance_t = \min(Rank_t - LowPerformance_t, 0.6)$$

$$HighPerformance = Rank_t - (LowPerformance_t + MiddlePerformance_t)$$

Using this specification the *LowPerformance* represents the bottom 20%, the *MiddlePerformance* represents the middle 60% and the *HighPerformance* represents the top 20%. Alternative specifications where the funds are divided into five equal groups are reported below.

I am particularly interested to explore whether there exists a marked difference among top performers. Therefore the cut-off point for the top quintile (20%) is substituted by 15%, 10% and 5%. If the flow performance relation differs substantially between top 10% and top 20% then this would indicate the importance of competition among the top performers, i.e. it is not just about being better than average but about winning the “competition”.

One possibility to estimate the model is treating each fund-year observation as an independent observation and using the OLS regression on the pooled dataset. However since these fund-year observations are suspicious to be dependent this might result in overstated t-statistics i.e. underestimate the standard errors of the coefficients. Following the Sirri and Tufano (1998), one alternative for this concern would be using the Fama and MacBeth (1973) two step procedure. The Fama and MacBeth two stages method is as follows: first estimate the flow-performance relationship separately for each year between 2004 and 2013 and then the final coefficients and t-statistics are estimated using the average of these coefficients estimates. The main advantage of this method is that it obtains more conservative results regarding the significance level of results. Therefore the Fama and MacBeth method is used as the main estimation method in this thesis.

## 5. Results and Discussion

### 5.1 Baseline Specification

The final sample includes 1292 U.S equity open end funds for the period between January 2004 and December 2014. The result from the main specification using Fama and MacBeth (1973) OLS method is reported in the table 4. In this specification fund's one year raw return is used as a measure of performance. To allow for nonlinearity of the relation a piecewise linear regression is used. Based on one year raw return funds are ranked and ordered. According to their ranking funds are divided into three performance group namely; *High*, *Middle* and *Low Performance*. The cut point for the *LowPerformance* is the bottom 20<sup>th</sup> percentile while the cut point for *High performance* varies from 5<sup>th</sup> percentile to 20<sup>th</sup> percentile. Other control variables are included in the model.

The results in table 4 in general confirm the convexity in the flow-performance relationship among the U.S mutual equity funds and indicate that this relation is not linear and it differs regarding the performance level of the fund. This result is in line with Sirri and Tufano (1998) Chevalier and Ellison (1995) and others findings. The coefficients in column "A" point out that there exist a marked difference between top and bottom performers and the convexity is economically significant. Moving from 80<sup>th</sup> to 85<sup>th</sup> percentile in raw return ranking would be awarded by 4.98% (0.997\*0.05) increase in fund's inflow. For the middle performance group there is a slightly positive relation between the flow and performance, although the results are less statistically significant (significant at 10%) compared to the results for the high performance group (significant at 1%). For the bottom performance group there is a negative relation however the results are not statistically significant.

The top quintile breakpoints have been changed from 20<sup>th</sup> percentile to 5<sup>th</sup> percentile. The columns B to D all have the same specifications as column A except for the high and mid performance ranking breakpoints. The column B separates the top 15<sup>th</sup> percentile performers from the rest, where columns C and D have breakpoints at 10<sup>th</sup> and 5<sup>th</sup> percentile respectively.

Comparing the regression through the column A to D suggest tightening the top performers group, results in more economically and statistically significant outcome. For instance, improving return ranking by 1 percentile in the top 20 percentile is awarded by 0.99% increase in inflow, while improving 1 percentile in the top 10<sup>th</sup> percentile is awarded by 2.67% increase in fund's inflow. This indicates that the competition among top performers to attract more inflow increases as the funds have higher ranks among top winners. To test whether this difference between HighPerformance quintiles at different breakpoints is statistically significant a Wald test is conducted. I can reject the null hypothesis that HighPerformance quintile at top 20 percentile and HighPerformance quintile at top 15 percentile are equal at 1% statistical significance. Similarly, I can reject the null hypothesis of equality of HighPerformance quintiles between column B and C as well as Column C and Column D, both at 1% statistical significance. Moreover, to test whether there is a statistically significant difference between different performance quintiles, a Wald test is conducted and the p-values for the null hypothesis that HighPerformance and LowPerformance are equal to each other are reported on the bottom line of the table for each regression. For each regression I can reject the null hypothesis that the HighPerformance are equal to MiddlePerformance and LowPerformance (p-values of 0.000) but I cannot reject the null hypothesis that MiddlePerformance and LowPerformance are equal to each other (p-values ranging from 0.014 to 0.138).

As we expected the investors are sensitive to fund's riskiness. The negative coefficient for Standard Deviation of Monthly return as a measure of riskiness indicates increasing the fund riskiness would result in decreasing fund's flow. Brown et al. (1996) find that fund managers, whom underperform the markets in the mid-year assessments, tend to increase the volatility of their fund in the second half year compared to those whom outperform the market. Chevalier and Ellison (1995) argue that the management fee structure in mutual fund industry provides an implicit incentive for fund manager to alter their portfolio riskiness. They provide some evidence that fund managers modify the riskiness of their portfolio accordingly with this incentive scheme. However, the negative relation between fund's volatility and inflow indicates that these incentives are not free of charge. For instance column A results indicate, 1% increase in the fund's volatility would be penalized by 0.8 % decrease in inflow. Ferreira, et. al (2012) study the relation between the level of convexity and the risk taking by the fund managers. They conclude

that as the level of convexity increase fund managers have more incentives to take higher level of risks.

The objective means flow variable (CategoryFlow) has a positive and statistically significant relation with flow indicating that the fund's flow is dependent on the average of inflow into the whole mutual fund industry. Other explanatory variables also have a significant impact on flow performance relation. The coefficient for fund age and size imply that smaller and younger funds are benefiting from greater flow than larger and older funds.

The positive statistically significant coefficient estimated for the Morningstar Rating is notable. The magnitude of the coefficient proposes that investors consider this rating in their investment allocation decision. The positive relation between rating and flow is in line with the findings from Guercio and Tkac (2008) who claim that mutual funds benefit from abnormal flow by investors response to the rating change of mutual funds.

Table 4: The flow performance relationship among U.S equity open end mutual fund 2004-2013

VARIABLES	(A)	(B)	(C)	(D)
LowPerformance <sub>t-1</sub>	-0.179 (0.144)	-0.196 (0.143)	-0.217 (0.142)	-0.264* (0.142)
MiddlePerformance <sub>t-1</sub>	0.0717* (0.0395)	0.0859** (0.0362)	0.102*** (0.0336)	0.130*** (0.0314)
HighPerformance <sub>t-1</sub>	0.997*** (0.239)	1.482*** (0.368)	2.674*** (0.686)	6.979*** (1.946)
SD <sub>t-1</sub>	-0.844** (0.364)	-0.853** (0.364)	-0.867** (0.363)	-0.884** (0.363)
CategoryFlow <sub>t</sub>	0.564*** (0.0845)	0.563*** (0.0845)	0.563*** (0.0845)	0.561*** (0.0845)
Log Age	-0.108*** (0.0108)	-0.108*** (0.0108)	-0.108*** (0.0108)	-0.108*** (0.0108)
Log TNA <sub>t-1</sub>	-0.0560*** (0.00547)	-0.0560*** (0.00547)	-0.0561*** (0.00548)	-0.0562*** (0.00549)
Total Fee <sub>t-1</sub>	0.0177 (0.0142)	0.0175 (0.0142)	0.0176 (0.0142)	0.0181 (0.0142)
Rating <sub>t-1</sub>	0.150*** (0.00936)	0.150*** (0.00936)	0.150*** (0.00935)	0.150*** (0.00937)
Constant	0.879*** (0.111)	0.881*** (0.111)	0.884*** (0.111)	0.887*** (0.112)
Number of Observations	8,635	8,635	8,635	8,635
R-squared	0.137	0.137	0.138	0.139
Wald Test (P- Values) HighPerformance=LowPerformance	0.0000	0.0000	0.0000	0.0002

The sample includes 1292 open end U.S equity mutual fund. For the one year raw return analysis there are 8635 fund-year observations in the sample. The table shows the OLS coefficient estimate using the growth rate of fund new inflow as dependent variable. Independent variables include fractional performance rank and other control variables. Here the performance measures by fund's one year raw return. The performance group represents the fund return ranking relative to other funds with same prospectus objective category and in same year which range from zero to one. The piecewise linear segments in the Column "A" is defined as: LowPerformance<sub>t</sub> = min (Rank<sub>t</sub>, 0.2), MiddlePerformance<sub>t</sub> = min (Rank<sub>t</sub> - LowPerformance<sub>t</sub>, 0.6) and HighPerformance<sub>t</sub> = Rank<sub>t</sub> - LowPerformance<sub>t</sub> - MiddlePerformance<sub>t</sub>. This procedure results in a cut point of 20% for the HighPerformance group. Furthermore, in the columns "B" to "D" the cut points for the HighPerformance funds are substituted by 15, 10 and 5 percent respectively. Control variables include fund riskiness measures by standard deviation of fund monthly return over past 12 month return (SD), Average inflow into the fund with same objective category (CategoryFlow), Natural Logarithm of fund age lagged by one year (Log Age), size of fund measured by Natural logarithm of fund's total net asset (Log TNA), Total fee as a percentage of asset under management which is calculated by summing up the annual net expense ratio and upfront load fee amortized over seven years period (Total Fee) and finally the Morningstar rating lagged by one year. The p-values of the Wald test are reported on the bottom line for each regression. The null hypothesis is if the HighPerformance and LowPerformance coefficients are equal. The standard errors and t-statistics are calculated using Fama and MacBeth (1973) two steps procedure method. P-values are presented in the parentheses below the coefficients.

\*\*\* 1% Significance level

\*\* 5% Significance level

\* 10% Significance level

## 5.2 Alternative Specification

In this section analyses are repeated using different measures of riskiness and performance. As an alternative performance measures the one year raw return is substituted by two, three and five years return. Moreover, previous literature suggests that fees charged by mutual fund affect the convexity of the flow-performance relationship. The fund's raw return is substituted by the fund's return net of management expenses and fees. Accordingly, standard deviation of monthly return over two, three and five years period is calculated and controlled in the model as a measure of riskiness. For each of these time periods the flow-performance relationship is analyzed using two different breakpoints for the top performer group, first for top 20<sup>th</sup> percentile (table 5 columns A, C, E and G) and then for top 10<sup>th</sup> percentile (table 5 columns B, D, F and H). Columns A and B rank funds based on their two years raw return, columns C and D rank them based on three years raw return, columns E and F rank funds based on their five year raw return. Columns G and H rank them based on one year return net of expenses and finally, column I rank them on the basis of five quintile groups instead of three quintiles. Using the similar methodology the piecewise segments for five quintiles are defined as follows:

$$Low_t = \min(Rank_t, 0.2)$$

$$Fourth_t = \min(Rank_t - Low_t, 0.2)$$

$$Third_t = \min(Rank_t - Low_t - Fourth_t, 0.2)$$

$$Second_t = \min(Rank_t - Low_t - Fourth_t - Third_t, 0.2)$$

$$\text{And } High_t = Rank_t - (Low_t + Fourth_t + Third_t + Second_t)$$

Table 5: Flow performance relationship, alternative specification

VARIABLES	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)
LowPerformance	0.273** (0.131)	0.152 (0.129)	0.232* (0.133)	0.152 (0.131)	0.0948 (0.143)	0.0323 (0.141)	-0.169 (0.145)	-0.214 (0.145)	
MiddlePerformance	0.0314 (0.0418)	0.114*** (0.0360)	0.0478 (0.0413)	0.124*** (0.0361)	0.192*** (0.0550)	0.247*** (0.0445)	0.0633* (0.0384)	0.0982*** (0.0332)	
HighPerformance	1.658*** (0.269)	3.872*** (0.784)	2.051*** (0.298)	5.506*** (0.906)	1.633*** (0.333)	4.071*** (0.892)	1.039*** (0.240)	2.732*** (0.698)	
Low $t_{-1}$									-0.0673 (0.164)
Fourth $t_{-1}$									-0.0782 (0.131)
Third $t_{-1}$									0.147 (0.138)
Second $t_{-1}$									0.107 (0.167)
High									0.949*** (0.281)
SD $t_{-1}$	-0.143 (0.465)	-0.144 (0.467)	-0.527* (0.320)	-0.571* (0.318)	-2.626*** (1.012)	-2.630*** (1.015)	-0.843** (0.366)	-0.863** (0.365)	-0.847** (0.364)
CategoryFlow $t$	0.625*** (0.101)	0.624*** (0.101)	0.678*** (0.127)	0.674*** (0.126)	0.601*** (0.156)	0.599*** (0.156)	0.574*** (0.0848)	0.573*** (0.0849)	0.564*** (0.0845)
Log Age $t_{-1}$	-0.104*** (0.0110)	-0.103*** (0.0109)	-0.0938*** (0.0112)	-0.0945*** (0.0111)	-0.0553*** (0.0135)	-0.0545*** (0.0134)	-0.107*** (0.0109)	-0.107*** (0.0108)	-0.108*** (0.0108)
Log TNA $t_{-1}$	-0.0505*** (0.00554)	-0.0505*** (0.00554)	-0.0486*** (0.00566)	-0.0484*** (0.00563)	-0.0499*** (0.00640)	-0.0498*** (0.00637)	-0.0579*** (0.00524)	-0.0580*** (0.00524)	-0.0559*** (0.00547)
Total Fee $t_{-1}$	0.0126 (0.0140)	0.0122 (0.0140)	-0.00135 (0.0145)	-0.00283 (0.0144)	-0.0112 (0.0159)	-0.0103 (0.0157)	- (-)	- (-)	0.0176 (0.0142)
Rating $t_{-1}$	0.121** (0.00913)	0.121** (0.00915)	0.0916*** (0.00965)	0.0886*** (0.00930)	0.0834*** (0.0139)	0.0822*** (0.0137)	0.146*** (0.00922)	0.146*** (0.00922)	0.150*** (0.00935)
Constant	0.732*** (0.109)	0.738*** (0.109)	0.785*** (0.112)	0.799*** (0.112)	0.833*** (0.123)	0.836*** (0.123)	0.948*** (0.0964)	0.953*** (0.0964)	0.870*** (0.112)
Number of Observations	7,869	7,869	7,044	7,044	4,811	4,811	8,645	8,645	8,635
R-squared	0.136	0.137	0.136	0.142	0.126	0.129	0.137	0.138	0.137
Wald Test (P- Values)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0017
HighPerformance=LowPerformance									

The table shows the OLS coefficient estimate using the growth rate of fund new inflow as dependent variable. Independent variables include fractional performance rank and other control variables. The performance rank calculated based on fund return over different time periods. Columns A and B rank funds based on their two years raw return, column C and D ranks them based on fund's three years raw return, column E and F ranks funds based on their five years raw return and finally column G and H ranks them based on one year return net of total fee. The piecewise linear segments for columns A to H are defined similar to the previous section. Columns A, C, E and G have the breakpoints for the HighPerformance group at the 20<sup>th</sup> percentile where the cut points for the Column B, D, F and H are at 10<sup>th</sup> percentile. In Column I the piecewise linear segments divide the funds into five equal quintiles and is defined as  $Low_t = \min(Rank_t, 0.2)$ ,  $Fourth_t = \min(Rank_t - Low_t, 0.2)$  and so forth. Control variables include: Average inflow into the fund with same objective category (CategoryFlow), Natural Logarithm of fund age lagged by one year (Log Age), size of fund measured by Natural logarithm of fund's total net asset (Log TNA), Total fee as a percentage of asset under management which is calculated by summing up the annual net expense ratio and upfront load fee amortized over seven years period (Total Fee), the Morningstar rating lagged by one year and the finally fund riskiness measures by standard deviation of fund monthly return, here the standard deviation calculated consistently to the time interval of the return as measure of performance. For instance for column A and B where performance is measured by two year raw return the standard deviation of the fund monthly return over past 24 month return is calculated and so forth through column I. The p-value of the Wald test is reported on the bottom line for each regression. The null hypothesis is if the HighPerformance and LowPerformance coefficients are equal. The standard errors and t-statistics are calculated using Fama and MacBeth (1973) two steps procedure method. P-values are presented in the parentheses below the coefficients.

\*\*\* 1% Significance level

\*\* 5% Significance level

\* 10% Significance level

The results in table 5 are consistent with my findings in previous section. However, by using the ranking based on two, three and five year performance, the existence of convexity can be observed for the middle and high performance quintiles. Surprisingly, in the two and three year analyzes the LowPerformance quintile is positive and statistically significant. The minor positive relation for middle performers and significant positive relation for high performers still exist. Moreover, there is a difference among the top performers irrespective of the performance measures that is used. The difference is both statistically and economically significant. For instance when we compare the results from column A and B, improving the performance ranking by 1 percentile in top 20<sup>th</sup> percentile group is rewarded by 1.65% increase in the flow while improving the performance ranking by 1 percentile in top 10<sup>th</sup> percentile is rewarded by 3.87% increase in the flow. The similar difference can be observed in the columns C and D as well as columns E and F. Analyzing and comparing the flow performance relation using three and five years return as a measure of performance results in the similar outcome. Using the net return as a measure of performance, (columns G and H) results in similar outcomes. Convexity still exists and the flow performance relationship becomes sharper as the funds are in the higher ranking quintile. While improving fund's net return rank by 1 percentile relative to other funds within top 20 percentile increase the flow by 1.04%, improving fund's net return rank by 1 percentile within top 10 percentile increase flow by 2.7%. To check whether the results in columns A, C, E, and G are statistically different from columns B, D, F and H, respectively, I use the Suest command and conduct a Wald test separately for each regression. I can reject the null hypothesis that results in columns A, C, E and G are equal to results in columns B, D, F and H, respectively.

In column I the five quintiles piecewise segment division is used and the convexity in the flow performance relationship can be observed. The results are only statistically significant for the top quintile. Using a Wald test I can reject the null hypothesis that the High quintile is equal to other four quintile groups (p-values ranging from 0.0017 to 0.0032) but I cannot reject the null hypothesis that other four quintiles are equal to one another (p-values ranging from 0.056 to 0.142). The result in column I once again suggests that performance of the funds in top quintile is rewarded intensively i.e. the fund can benefit markedly from flow performance relationship by not only performing better than average but also by securing their position in top performance quintile. The standard deviation of fund's return is not statistically significant in columns A and B but statistically significant in the other columns. The negative coefficient indicates that investors are sensitive to fund's riskiness regardless of the riskiness measures. Average flow into



the whole mutual fund industry explains part of the flow relation i.e. the individual fund flows are affected by the flows into the fund with same objective category. Once again the coefficients on the fund age indicate that younger funds benefit more from the flow-performance relationship. Furthermore, the size coefficients (Log TNA) show the smaller funds can gain more flow. Finally, a Wald test is again conducted to test the convexity of the flow performance relationship. The null hypothesis that HighPerformance and LowPerformance are equal can be rejected (p-values of 0.000). Also the null hypothesis that HighPerformance and MiddlePerformance are equal can be rejected (p-values ranging from 0.000 to 0.002) however, once again I cannot reject the null hypothesis that MiddlePerformance and LowPerformance quintiles are equal (p-values ranging from 0.131 to 0.852).

The results in table 5 create a steady image from flow performance relationship in mutual fund industry. There is convex relation between fund return and flow. However, by using the longer time periods of historical return the negative relation for low performers does not exist anymore. Still, there is slightly positive relation for middle performers and a strong positive relation for high performers. The results for both middle and high performers are statistically significant. Comparing the result among top performers using different cut-points provide evidences that the competition is tougher among these top performers and the funds can gain additional marked inflow if they can secure their position among top 10% funds.

### **5.3 Origin of convexity**

In this section I aim to explore the root of the convex flow-performance relationship observed in the previous sections and to understand better which factors might lead to a convex flow-performance relationship. To this aim the sample is broken into subsamples based on the funds age and types and the relation is estimated separately for these different groups.

The literature talks about differences in flow-performance relationship among different age categories. To test this hypothesis, I first divide the funds based on their age into two subsamples of young and old funds. All the funds with an age between 2 and 5 at the point of observation are named young funds and the funds with age more than 5 years are called as the old funds. Then, the flow-performance relationship is estimated separately for these two groups. More than 85%

of fund-year observations in the datasets are old funds. The results from the regressions are reported in table 6 column A and B for young and old funds, respectively. The convexity of the flow performance relationship can be observed in both groups. However, the flow of young funds seems to be more sensitive to the historical performance (one year raw return) than the old funds. I test the hypothesis that convexity is equal between young and old funds to confirm that there exists a difference in reaction of investors to old and young funds' performance. I can reject this hypothesis. For the lowPerformance quintile there is a noticeable differences in the outflow of the young funds compared to the old funds however the results are not statistically significant. Furthermore, for the middle and high performance quintiles flow increase more for the young funds compared to older funds. One possible explanation for this could be the fact that investors are more sensitive to the young funds performance due to the lack of historical performance of these funds. Another explanation is that since most of these younger funds are also smaller funds compare to older funds, increasing the same amount of flow in dollar value means a higher flow in percentage for them. These results are in line with Chevalier and Ellison (1997) findings although they use a semiparametric model to estimate and compare the flow performance relationship between young and old funds.

Furthermore, the literature also notices the impact of fund type on the flow performance relationship (for instance Keswani & Stolin, (2012)). My dataset includes information about whether the funds are designed for institutional or non-institutional investors. Based on this information I divide the funds and analyze the relation separately for them. The results are presented in the table 6 column C and D for the institutional and non-institutional funds, respectively. A convex flow-performance relationship can be observed for the non-institutional funds similar to the relation that exists for the aggregate sample. However, the performance sensitivity for the non-institutional top quintile performance is higher than for the full sample. For the institutional funds the flow performance relationship is statistically insignificant. The size of estimated coefficients for the HighPerformance quintiles is close to the estimated coefficient for the LowPerformance quintile. This could be interpreted as institutional investors sell past loser almost at the rate that they buy past winners. Keswani & Stolin, (2012) show that different types of investors generate significant differences in the reaction to the fund performance. They conclude that flow-performance convexity in mutual funds is mainly driven by the reaction of retail investors rather than institutional investors. This can be confirmed by my result which

indicates that huge bonuses for the past winners are related to non-institutional funds i.e. retail investors.

Table 6: Flow performance relationship among certain subgroups

VARIABLES	Panel 1		Panel 2	
	(A)	(B)	(C)	(D)
LowPerformance <sub>t-1</sub>	-0.914 (0.610)	-0.142 (0.137)	-0.590 (0.383)	-0.174 (0.151)
MiddlePerformance <sub>t-1</sub>	0.303** (0.151)	0.0690** (0.0312)	0.0851 (0.0852)	0.104*** (0.0361)
HighPerformance <sub>t-1</sub>	5.316* (2.819)	2.051*** (0.575)	0.504 (1.174)	3.334*** (0.814)
SD <sub>t-1</sub>	1.114 (1.417)	-1.285*** (0.325)	-0.437 (0.923)	-0.943** (0.388)
CategoryFlow <sub>t</sub>	1.634*** (0.308)	0.365*** (0.0782)	0.249 (0.221)	0.647*** (0.0910)
Log Age <sub>t-1</sub>	-0.214** (0.0965)	-0.0641*** (0.0109)	-0.135*** (0.0286)	-0.0980*** (0.0116)
Log TNA <sub>t-1</sub>	-0.106*** (0.0200)	-0.0448*** (0.00479)	-0.0551*** (0.0155)	-0.0591*** (0.00583)
Total Fee <sub>t-1</sub>	-0.000473 (0.0514)	0.0191 (0.0129)	-0.105 (0.0719)	0.0185 (0.0156)
Rating <sub>t-1</sub>	0.215*** (0.0300)	0.132*** (0.00857)	0.174*** (0.0236)	0.145*** (0.0102)
Constant	1.575*** (0.427)	0.636*** (0.0975)	1.055*** (0.323)	0.911*** (0.121)
Number of Observations	1,240	7,395	1,797	6,838
R-squared	0.143	0.105	0.112	0.149
Wald Test (P- Values)	0.0000	0.0000	0.0000	0.0000
HighPerformance=LowPerformance				

The table presents the flow performance relationship among different subsamples. The OLS coefficients are estimated using the growth rate of fund new inflow as dependent variable. Independent variables include fractional performance rank and other control variables. The performance rank calculated based on fund one year raw return. Panel 1 compares the flow performance relationship between young and old funds. Young funds defined as the funds with age between 2 and 5 years and the old funds subsample consists of funds with age older than 5 years column A and B show the young and old funds estimation respectively. Panel 2 compares the flow-performance relationship between institutional and non-institutional funds. Column C and D shows the coefficients estimation for institutional and non-institutional subsamples respectively. Control variables include: Average inflow into the fund with same objective category (CategoryFlow), Natural Logarithm of fund age lagged by one year (Log Age), size of fund measured by Natural logarithm of fund's total net asset (Log TNA), Total fee as a percentage of asset under management which is calculated by summing up the annual net expense ratio and upfront load fee amortized over seven years period (Total Fee), the Morningstar rating lagged by one year and finally fund riskiness measures by standard deviation of fund monthly return. The p-values of the Wald test is reported on the bottom line for each regression. The null hypothesis is if the HighPerformance and LowPerformance coefficients are equal. The standard errors and t-statistics are calculated using Fama and MacBeth (1973) two steps procedure method. P-values are presented in the parentheses below the coefficients.

\*\*\* 1% Significance level

\*\* 5% Significance level

\* 10% Significance level

## 6. Summary and Conclusion

The main objective of this thesis was to explore whether there is a difference in the flow-performance relationship among the best performing funds. To my knowledge, most of the literature studies flow-performance relationship using different models and different set of control variables to estimate the shape of this relationship in general and there is no study that focuses on the flow-performance relationship among top performing funds. To this end, following Sirri and Tufano (1998) and others I first explore the flow performance relationship among the U.S open end equity mutual funds for the time interval between January 2004 and December 2013. The results suggest that there exists a convex relation between flow and performance of U.S mutual funds. For the middle performance group of funds there is a slightly positive relation while high performing are rewarded hugely in terms of increasing the inflow. Sirri and Tufano (1998) report the same type of convexity. I extend their analysis of flow performance by examining the relation among top performers. Particularly, I was interested to test whether the funds can obtain additional marked inflow by winning the competition against their peers.

My results suggest that the level of convexity within the top 20 percentile funds increases with fund performance. The level of convexity among the top 20 percentile is beyond what I observe in the flow performance relation across whole sample. I find that there exists a marked difference between being among top 20<sup>th</sup> percentile and being among top 10<sup>th</sup> percentile and the difference is significant. Regardless of performance measurement the funds within top 10<sup>th</sup> percentile can obtain a flow which is more than twice as much as funds within top 20<sup>th</sup> percentile can obtain.

Using longer time periods for historical return improves the results in terms of statistical significance. In general the results suggest that convexity in the flow-performance relationship exist and there is a marked difference among top performers regardless of return measurements. I also document a positive relation between Morningstar rating and fund flow.

The results also suggest that there is a negative relation between fund's portfolio riskiness and its flow. Chevalier & Ellison (1997) conclude that the management fee structures in the mutual fund industry create an incentive for fund managers to take higher level of risk in order to gain flow. This could form a conflict between fund managers and investors since the managers might

increase their portfolio riskiness to improve their performance. The negative relation between fund riskiness and flow notice that increasing fund's riskiness is not free of charge since the investors are sensitive to the fund volatility.

Finally, I examine the relation separately for the funds that are younger and older than 5 years. The results suggest that investors are more sensitive to the young funds performance and response to the best and the worst fund's performance more intensely than they response to the old fund's performance. Furthermore, I estimate the model for institutional and non-institutional funds and results propose that the main part of convexity in the relationship comes from non-institutional investors rather than institutional investors. Non-institutional investors invest disproportionately in the best past performers while they fail to outflow from the worst past performers. However, for the institutional investors there is no statistically significant relation between fund flow and performance.

This thesis aimed to answer the question whether competition in mutual fund industry to attract higher inflow is to perform better than average or if it is about winning the competition by achieving the highest possible position in the market. I find funds that have a position within top 10% are rewarded much more strongly than the next 10% in terms of inflow. My conclusion is that the competition in the mutual fund industry is not about performing better than average, rather it is about winning the competition and securing a position among top 10%.

My results have implication both for the fund managers and regulators. For the fund managers, the results suggest that the main competition for attracting the higher investment is among the best performers. The investors invest highly disproportionately to the best past performers and the level of this disproportionality increase by the level of fund performance. This creates an opportunity for fund managers to achieve additional marked inflow by securing their position among top percentile performers. I conclude if fund managers aim to attract the highest possible investment they need to secure their position among the top 10% in terms of performance.

Moreover, considering contradictory evidence about fund's performance persistence, the way that investors respond to top fund's performance can be interpreted as an irrational behaviour. This should encourage regulators to provide facilities for investors, where they can improve their investment knowledge.

Access to the fund's data has been gradually increased over past decades and so many researchers attempt to understand the behavior of investors and fund managers. However, there are still a lot of room to learn more about investors' behavior, intermediaries' contribution and fund managers. One of the potential future areas to study in the fund industry would be focusing on fund's skewness i.e. to investigate if return skewness would affect flow-performance relationship for top quintiles performers.

## References:

- morningstar.com*. (2014). Retrieved 2014, from Morning Star Glossary
- Alves, C., & Mendes, V. (2011). Does performance explain mutual fund flows in small markets? The case of Portugal. *Portuguese Economic Journal*, 129-147.
- Ballester, C. P. (2014). Determinants of Equity Pension Plan Flows. *Estudios de Economia*, 125-148.
- Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 1269–1295.
- Brown, K. C., Harlow, W. V., & Starks, L. T. (1996). Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry. *The Journal of Finance*, 85-110.
- Brown, G., Draper, P., & E. M. (1997). Consistency of UK Pension Fund Investment Performance. *Journal of Business Finance & Accounting*, 155-178.
- Brown, S. J., Goetzman, W., Ibbotson, R. J., & Ross, S. A. (1992). Survivorship bias in performance studies. *Review of financial studies*, 553-580.
- Brown, S., & Goetzmann, W. (1995). Attrition and mutual fund performance. *Journal of Finance*, 679–698.
- Carhart, M. M. (1997). On persistence of mutual fund performance. *The Journal of Finance*, 57-82.
- Chevalier, J., & Ellison, G. (1997). Risk Taking by Mutual Funds as a Response to Incentives. *Journal of Political Economy*, 1167-1200.
- Dahlquist, M., Engström, S., & Söderlind, P. (2000). Performance and Characteristics of Swedish Mutual Funds. *Journal of Financial and Quantitative Analysis*, 409–423.
- Del Guercio, D., & Tkac, P. A. (2002). The determinants of the flow of funds of managed portfolios: mutual funds versus pension funds. *Journal of Financial and*, 523–558.
- Fama, E., & MacBeth, J. (1973). Risk, return and equilibrium: Empirical tests. *Journal of Political Economy*, 607–636.
- Ferreira, M. A., Keswani, A., Miguel, A. F., & Ramos, S. B. (2012). The flow-performance relationship around the world. *Journal of Banking & Finance*, 1759–1780.
- Goetzmann, W., & Peles, N. (1997). Cognitive dissonance and mutual fund investors. *Journal of Financial Research*, 145-158.
- Grinblatt, M., & Titman, S. (1992). The persistence of mutual fund performance. *Journal of Finance*, 1977–1984.
- Guercio, D. D., & Tkac, P. A. (2008). Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow. *JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS*, 907–936.

Miri / Competition in the American Mutual Fund Industry: An Empirical Study

Hendricks, D., Patel, J., & Zeckhauser, R. (1993). Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance. *Journal of Finance*, 93–130.

Huang, J., Sialm, C., & Zhang, H. (2011). Risk shifting and mutual fund performance. *Review of Financial Studies*, 2575–2616.

Ibbotson, R., & Goetzmann, W. (1994). Do winners repeat? *Journal of Portfolio Management*, 9–18.

ICI. (2014). *Mutual Fund Fact Book: Industry Trends and Statistics*. Washington, D.C: Investment Company Institute.

Ippolito, R. A. (1992). Consumer reaction to measures of poor quality: Evidence from the mutual fund industry. *Journal of Law and Economics*, 45–70.

Jain, P. C., & Wu, J. S. (2000). Truth in Mutual Fund Advertising: Evidence on Future Performance and Fund Flows. *The Journal of Finance*, 937-958 .

Kempf, A., & Ruenzi, S. (2008). Family Matters: Rankings Within Fund Families and Fund Inflows. *Journal of Business Finance & Accounting*, 177–199.

Keswani, A., & Stolin, D. (2012). Investor reaction to mutual fund performance: evidence from UK distribution channels. *The Journal of Financial Research*, 425–450.

Lynch, A. W., & Musto, D. K. (2003). How investors interpret past fund returns. *Journal of Finance*, 2033–2058.

Malkiel, B. G. (1995). Returns from investing in equity mutual funds. *Journal of Finance*, 549–573.

McDonald, J. G. (1974). Objectives and Performance of Mutual Funds, 1960–1969. *Journal of Financial and Quantitative Analysis* , 311-333.

Navone, M., Pagani, M., & Pantos, T. D. (n.d.). The Determinants of the Convexity in the Flow-Performance Relationship.

Sapp, T., & Tiwari, A. (2004). Does stock return momentum explain the Smart Money effect? *Journal of Finance*, 2605-2622.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *Journal of finance*, 1589-1622.

Smith, K. V. (1978). Is fund growth related to fund performance? *Portfolio Management*, 49–54.

Spitz, A. E. (1970). Mutual Fund Performance and Cash Inflows. *Applied Economics* 2, 141–145.



Databases

Morningstar. (2014) *Morningstar Direct*. (Online). Available at: Subscription Service (Accessed: February 2014)

Bloomberg. (2014) *Bloomberg*. (Online). Available at: Subscription Service. (Accessed: February 2014)